



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :		A2	(11) International Publication Number: WO 98/39446
C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17			(43) International Publication Date: 11 September 1998 (11.09.98)
(21) International Application Number: PCT/US98/04482			MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).
(22) International Filing Date: 6 March 1998 (06.03.98)			
(30) Priority Data:			
60/040,162	7 March 1997 (07.03.97)	US	
60/040,333	7 March 1997 (07.03.97)	US	
60/038,621	7 March 1997 (07.03.97)	US	
60/040,161	7 March 1997 (07.03.97)	US	
60/040,626	7 March 1997 (07.03.97)	US	
60/040,334	7 March 1997 (07.03.97)	US	
60/040,336	7 March 1997 (07.03.97)	US	
60/040,163	7 March 1997 (07.03.97)	US	
60/043,580	11 April 1997 (11.04.97)	US	
60/043,568	11 April 1997 (11.04.97)	US	
<i>(Continued on the following page)</i>			
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOP-PET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BED-NARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda,			
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).			
<b>Published</b>			
Without international search report and to be republished upon receipt of that report.			
With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.			
Date of receipt by the International Bureau: 06 April 1998 (06.04.1998)			

(54) Title: 70 HUMAN SECRETED PROTEINS

## (57) Abstract

The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

60/043,314	11 April 1997 (11.04.97)	US	60/047,598	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US
60/043,569	11 April 1997 (11.04.97)	US	60/047,613	23 May 1997 (23.05.97)	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,311	11 April 1997 (11.04.97)	US	60/047,582	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,671	11 April 1997 (11.04.97)	US	60/047,596	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,674	11 April 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,669	11 April 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,312	11 April 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05.97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,313	11 April 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05.97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,672	11 April 1997 (11.04.97)	US	60/047,599	23 May 1997 (23.05.97)	US	60/056,636	22 August 1997 (22.08.97)	US
60/043,315	11 April 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05.97)	US	60/056,874	22 August 1997 (22.08.97)	US
60/043,578	11 April 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05.97)	US	60/056,910	22 August 1997 (22.08.97)	US
60/043,576	11 April 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08.97)	US
60/043,670	11 April 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05.97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,594	23 May 1997 (23.05.97)	US	60/056,845	22 August 1997 (22.08.97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,589	23 May 1997 (23.05.97)	US	60/056,892	22 August 1997 (22.08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05.97)	US	60/056,632	22 August 1997 (22.08.97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,876	22 August 1997 (22.08.97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,881	22 August 1997 (22.08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,909	22 August 1997 (22.08.97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,875	22 August 1997 (22.08.97)	US
60/047,503	23 May 1997 (23.05.97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,862	22 August 1997 (22.08.97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22.08.97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22.08.97)	US
60/047,584	23 May 1997 (23.05.97)	US	60/056,630	22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
60/047,500	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
60/047,587	23 May 1997 (23.05.97)	US	60/056,662	22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047,492	23 May 1997 (23.05.97)	US	60/056,872	22 August 1997 (22.08.97)	US			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## 70 Human Secreted Proteins

### *Field of the Invention*

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

### ***Summary of the Invention***

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, 5 and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

10

### ***Detailed Description***

#### **Definitions**

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original 15 environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

20 In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce 25 a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid 30 sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated 35 amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 12301 Park Lawn Drive, Rockville, Maryland 20852, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5       The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and 10 double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability 15 or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

20       The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, 25 as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or 30 without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a 35 nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenylation, sulfation, transfer-RNA mediated addition of amino acids to proteins

- 5 such as arginylation, and ubiquitination. (See, for instance, PROTEINS -  
STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W.  
H. Freeman and Company, New York (1993); POSTTRANSLATIONAL  
COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic  
Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990);  
10 Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y"  
refers to a polypeptide sequence, both sequences identified by an integer specified in  
Table 1.

- 15 "A polypeptide having biological activity" refers to polypeptides exhibiting  
activity similar, but not necessarily identical to, an activity of a polypeptide of the  
present invention, including mature forms, as measured in a particular biological assay,  
with or without dose dependency. In the case where dose dependency does exist, it  
need not be identical to that of the polypeptide, but rather substantially similar to the  
dose-dependence in a given activity as compared to the polypeptide of the present  
20 invention (i.e., the candidate polypeptide will exhibit greater activity or not more than  
about 25-fold less and, preferably, not more than about tenfold less activity, and most  
preferably, not more than about three-fold less activity relative to the polypeptide of the  
present invention.)

25 **Polynucleotides and Polypeptides of the Invention**

**FEATURES OF PROTEIN ENCODED BY GENE NO: 1**

- The translation product of Gene NO: 1 shares sequence homology with alpha-L-fucosidase which is thought to be important as a lysosomal enzyme that hydrolyzes  
30 fucose from fucoglycoconjugates. (See Accession No. gi/178409.) Lysosome  
fructosidase is involved in certain lysosome storage diseases. (See Biochem. Biophys.  
Res. Commun., 164(1):439-445 (1989).) Fucosidosis, an autosomal recessive  
lysosomal storage disorder characterized by progressive neurological deterioration and  
mental retardation. The disease results from deficient activity of alpha-L-fucosidase, a  
35 lysosomal enzyme that hydrolyzes fucose from fucoglycoconjugates. This gene likely  
encodes a novel fucosidase isoenzyme. Based on homology, it is likely that the  
translated product of this gene is also involved in lysosome catabolism of molecules and

that aberrations in the concentration and/or composition of this product may be causative in lysosome storage disorders. Preferred polypeptide fragments comprise the amino acid sequence PGHLLPHKWENC (SEQ ID NO: 257).

Gene NO: 1 is expressed primarily in stromal cells, and to a lesser extent in  
5 human fetal kidney and human tonsils.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fucosidosis and other lysosome storage disorders. Similarly,  
10 polypeptides and antibodies directed to the polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues of cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., stromal cells, kidney, tonsils, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue  
15 or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 1 to alpha-L-fucosidase  
20 indicates that polypeptides and polynucleotides corresponding to Gene NO: 1 are useful for the treatment of fucosidosis and general lysosomal disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
134 as residues: Met-1 to Leu-6, Thr-32 to Glu-39, Lys-80 to Lys-85, and Met-90 to  
Pro-96.

25

## FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of Gene No. 2 shares sequence homology with stromal cell-derived factor-2 (SDF-2) which is a novel secreted factor. See, for example, Gene, 176(1-2):211-214, (1996, Oct. 17.) The amino acid sequence of SDF-2 shows  
30 similarity to yeast dolichyl phosphate-D-mannose:protein mannosyltransferases, Pmt1p [Strahl-Bolsinger et al. Proc. Natl. Acad. Sci. USA 90, 8164-8168 (1993)] and Pmt2p [Lussier et al. J. Biol. Chem. 270, 2770-2775 (1995)], whose activities have not been detected in higher eukaryotes. Based on the sequence similarity, the translation product of this gene is expected to share certain biological activities with SDF-2, Pmt1p and  
35 Pmt2p.

Gene NO: 2 is expressed primarily in immune system tissue and cancerous tissues, such as liver hepatoma, human B-cell lymphoma, spleen in a patient suffering

from chronic lymphocytic leukemia, hemangiopericytoma, pharynx carcinoma, breast cancer, thyroid, bone marrow, osteoblasts and to a lesser extent in a few other tissues such as kidney pyramids.

Therefore, polynucleotides or polypeptides of the invention are useful as

- 5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the diseases and conditions which include, but are not limited to, disorders in kidney, liver, and immune organs, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell  
10 type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, liver, thyroid, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, spleen, B-cells, pharynx, thyroid, mammary tissue, bone marrow, osteoblasts and kidneys, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,  
15 urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 2 to stromal cell-derived

- 20 factor-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 2 are useful for diagnosis and therapeutic treatment of disorders in kidney, liver, and immune organs since stromal cells play important role in organ function. Stroma carries the blood supply and provides support for the growth of parenchymal cells and is therefore crucial to the growth of a neoplasm. Nucleic acids of the present invention  
25 comprise, but preferably do not consist of, and more preferably do not comprise, SEQ ID NO: 3 from US Patent No. 5,576,423, incorporated herein by reference, and shown herein as SEQ ID NO: 258).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

- 135 as residues: His-56 to Gly-65, Ala-74 to Ser-80, Ile-84 to Pro-97, Leu-124 to Glu-  
30 129, Glu-135 to Asp-143, Gly-175 to Ser-180, and Ala-194 to Thr-199.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 3

- The translation product of Gene NO: 3 shares sequence homology with LZIP-1, LZIP-2 and other leucine zipper proteins, which are thought to be important in nucleic acid binding. This gene has been reported in Mol. Cell. Biol. 17 (9), 5117-5126 (1997) as "Luman". Luman is a cyclic AMP response element (CRE)-binding protein/activating transcription factor 1 protein of the basic leucine zipper superfamily. It binds CREs in

vitro and activates CRE-containing promoters when transfected into COS7 cells. The complete amino acid sequence of Luman reported in Mol. Cell. Biol. 17 (9): 5117-5126 (1997) is:

MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALPLSEVPSDWEVDDLL  
5 CSLLSPPASLNILSSSNPCLVHHHDHTYSLPRETVSMDESESCKEGTQMTPQH  
MEELAEQEIAIRVLTDEEKSLLKEGLLPLTKTEEQILKRVRRKIRNKRSA  
QESRRKKKVYVGGLSERVLKYTAQNMELQNKVQLLEEQNLSLLDQLRKLQAM  
VIEISNKTSSSTCILVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPSE  
DPYQLELPALQSEVPKDSTHQWLGSDCVLQAPGNTSCLHYMPQAPS AEPPL  
10 EWPFPDLSS EPLCRGPILPLQANLTRKGWLPTGSPSVILQDRYSG (SEQ ID  
N:259).

Gene NO: 3 is expressed primarily in apoptotic T-cells and Soares senescent cells and to a lesser extent in multiple tissues and cell types, including, multiple sclerosis tissue, and hippocampus.

15 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. Similarly, polypeptides and antibodies directed to these  
20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and transplantation, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., multiple sclerosis tissue, hippocampus, bone marrow and cancerous and wounded  
25 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 3 to leucine zipper nucleic  
30 acid binding proteins indicates that polypeptides and polynucleotides corresponding to Gene NO: 3 are useful for diagnosis and treatment of immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. The secreted nucleic acid binding protein in the apoptotic tissues may be involved in the disposal of the DNA released by apoptotic cells. Furthermore, the studies conducted in support of Luman  
35 suggest that the translation product of this gene may be used to identify transcriptional regulation elements which in turn are useful in modulation of immune function.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 136 as residues: Asn-7 to Ser-12, Tyr-32 to Gly-38, Pro-55 to Tyr-60, Glu-70 to Thr-76, and Pro-104 to Leu-110.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of Gene NO: 4 shares sequence homology with a number of tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene, CO-029 tumor associated antigen protein, CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, which are thought to be important in development of cancer, immune system development and functions.

Gene NO: 4 is expressed primarily in cancers of several different tissues and to a lesser extent in normal tissue like prostate, skin and kidney.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers and disorders of the immune system, prostate and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, skin, prostate and immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., kidney, skin and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 4 to tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene, CO-029 tumor associated antigen protein, CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, indicates that polypeptides and polynucleotides corresponding to Gene NO: 4 are involved with the cellular control of growth and differentiation. Therefore, the translation product of this gene is believed to be useful for diagnosis and treatment of neoplasia and disorders of the kidney, skin and prostate. For example, recombinant protein can be produced in transformed host cells for diagnostic and prognostic applications. Alterations in the protein sequence are indicative of the presence of

malignant cancer, or of a predisposition to malignancy, in a subject. Gene therapy can be used to restore the wild-type gene product to a subject. Additionally, the antibodies are a useful tool for the identification of hematopoietic neoplasms, and may prove helpful for identifying morphologically poorly defined cells. Moreover, this protein can  
5 be used to isolate cognate receptors and ligands and identify potential agonists and antagonists using techniques known in the art. The protein also has immunomodulatory activity, regulates hematopoiesis and stimulates growth and regeneration as a male/female contraceptive, increases fertility depending on activin and inhibin like activities. Other uses are as a chemotactic agent for lymphocytes, treatment of  
10 coagulation disorders, an anti-inflammatory agent, an antimicrobial or analgesic and as a modulator of behavior and metabolism. The DNA can be used in genetic diagnosis or gene therapy, and for the production of recombinant protein. It can also be used to identify protein expressing cells, isolate related sequences, prepare primers for genetic fingerprinting and generate anti-protein or anti-DNA antibodies. In addition, residues 1-  
15 71, in the translation product for this gene are believed to be the extracellular domain. Thus, polypeptide comprising residues 1-71 or derivatives (including fragments) or analogs thereof, are useful as a soluble polypeptide which may be routinely used therapeutically to antagonize the activities of the receptor.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
20 137 as residues: Lys-118 to Phe-127, Asn-145 to Ala-160, and Thr-177 to Val-188.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 5

Gene NO: 5 is expressed primarily in human testes.  
Therefore, polynucleotides or polypeptides of the invention are useful as  
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the testes including cancer and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell  
30 type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the  
35 standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of Gene NO: 5 indicates that the protein product of this gene is useful for treatment/diagnosis of diseases of the testes, particularly testicular cancer since expression is observed primarily in the testes.

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
5 138 as residue: Gly-22 to Gln-30.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of Gene NO: 6 shares sequence homology with GALNS (N-acetylgalactosamine 6-sulphatase) which is thought to be important in the storage of  
10 the glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. See Genbank accession no. gil618426. Based on the sequence similarity, the translation product of this gene is expected to share biological activities with GALNS.

Gene NO: 6 is expressed primarily in human bone marrow.  
Therefore, polynucleotides or polypeptides of the invention are useful as  
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate, e.g., Morquio A syndrome. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential  
20 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly involving cell storage disorder, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from  
25 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 6 to N-acetylgalactosamine 6-sulphatase indicates that polypeptides and polynucleotides corresponding to Gene  
30 NO: 6 are useful for the treatment and diagnosis of storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. Such disorders are known in the art and include, e.g., Morquio A syndrome which is caused by an error of mucopolysaccharide metabolism with excretion of keratan sulfate in urine. Morquio A syndrome is characterized by severe skeletal defects with short stature, severe deformity  
35 of spine and thorax, long bones with irregular epiphyses but with shafts of normal length, enlarged joints, flaccid ligaments, and waddling gait; autosomal recessive inheritance.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 139 as residues: Gly-29 to Pro-36 and Glu-57 to Leu-64.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 7

5 The translation product of Gene NO: 7 shares sequence homology with carboxy peptidase E and H (carboxypeptidase E is thought to be important in the biosynthesis of numerous peptide hormones and neurotransmitters). The translation product of this gene also shares sequence homology with bone-related carboxypeptidase "OSF-5" from the mouse. See European patent application EP-588118-A. Based on the sequence 10 similarity to OSF-5, the translation product of this gene will hereinafter sometimes be referred to as "human-OSF-5" or "hOSF-5".

Gene NO: 7 is expressed primarily in tumor cell lines derived from connective tissues including chondrosarcoma, synovial sarcoma, Wilm's tumor and rhabdomyosarcoma and to a lesser extent in a myeloid progenitor cell line, bone 15 marrow, and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various cancers involving the skeletal system and connective tissues in 20 general, in particular at cartilage interfaces. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system and various other tumor tissues, expression of this gene at significantly higher or lower levels may routinely be 25 detected in certain tissues (e.g., connective tissues and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The restricted tissue distribution and homology of Gene NO: 7 to carboxypeptidase E and mouse OSF-5 indicates that polypeptides and polynucleotides corresponding to Gene NO: 7 are for processing of peptides to their mature form that may have various activities similar to the activities of neuropeptides but in the periphery. In addition the abundance of expression in cancer tissues indicates that 35 aberrant expression and subsequent processing may play a role in the progression of malignancies, e.g., growth factor and/or adhesion factor activities. In particular, the expression of this gene is restricted to connective tissues and embryonic tissues.

Furthermore, it is overexpressed in cancers of these same tissues (i.e., in sarcomas). Moreover, hOSF-5 shares very strong sequence similarity with mOSF-5 which is a known bone growth factor and is thought to be useful in obtaining products for the diagnosis and treatment of bone metabolic diseases, e.g., osteoporosis and Paget's disease. Like OSF-5, the translation product of this gene is believed to be a bone-specific carboxypeptidase which acts as an adhesion molecule/growth factor and takes part in osteogenesis at the site of bone induction. hOSF-5 can, therefore, be used to treat bone metabolic diseases, osteoporosis, Paget's disease, osteomalacia, hyperostosis or osteopetrosis. Furthermore, hOSF-5 can be used to stimulate the regeneration of bone at the site of mechanical damage, e.g., accidentally or surgically caused fractures.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 140 as residues: Leu-24 to Val-30, Ala-89 to Lys-94, Phe-150 to Trp-157, Leu-162 to Asp-167, Asp-187 to Ser-199, His-241 to Asp-254, and Pro-362 to Asp-376.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 8**

Gene NO: 8 is expressed primarily in bone marrow, and to a lesser extent in an erythroleukemia cell line.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematological disorders including cancer and anemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematologic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 8 are useful as a growth factor for hematopoietic stem cells or progenitor cells, e.g., in the treatment of bone marrow stem cell loss in chemotherapy patients and in the treatment of kidney disease.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 141 as residues: Gly-30 to Lys-35.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 9**

Gene NO: 9 is expressed primarily in neutrophils.

Therefore, polynucleotides or polypeptides of the invention are useful as

- 5 reagents for differential identification of the cell type present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the cell type indicated. For a number of disorders of the above tissues or cells,
- 10 particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., neutrophils, bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
- 15 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 9 are useful for immune modulation or as a growth factor to stimulate neutrophil differentiation or proliferation that may be useful in the treatment of neutropenia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 142 as residues: Thr-22 to Pro-37.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 10**

- 25 Gene NO: 10 is expressed primarily in the epidermis.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the epidermis such as psoriasis or eczema or may be involved in the normal proliferation or differentiation of the epithelial cells or fibroblasts constituting the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 10 are useful for diagnosis and treatment of skin conditions and as an aid in the healing of various epidermal injuries including wounds, and diabetic ulcers.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 143 as residues: Ser-3 to Ser-9 and Trp-27 to Glu-32.

## 10 FEATURES OF PROTEIN ENCODED BY GENE NO: 11

The translation product of Gene NO: 11 shares sequence homology with phosphatidylcholine 2-acylhydrolase (PLA2). See, for example, Genbank accession no. gi|190004. PLA2 is involved in inflammation, where it is responsible for the conversion of cell membrane phospholipids into arachidonic acid. Arachidonic acid in turn feeds into both the lipoxygenase and cyclooxygenase pathways to produce leukotrienes (involved in chemotaxis, vasoconstriction, bronchoconstriction, and increased vascular permeability) and prostaglandins (responsible for vasodilation, potentiate edema, and increased pain). Diseases in which PLA2 is implicated as a major factor include rheumatoid arthritis, sepsis, ischemia, and thrombosis. The inventors refer to the translation product of this gene as PLA2-like protein based on the sequence similarity. Furthermore, owing to the sequence similarity PLA2 and PLA2-like protein are expected to share certain biological activities.

Gene NO: 11 is expressed primarily in human cerebellum and in T-cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cerebellum and Purkinje cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, bone marrow, T-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 11 are useful for diagnosis and treatment of cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. This gene is also useful as a chromosome marker. It is believed to map to Chr.15, D15S118-D15S123.

5

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 12

Gene NO: 12 is expressed primarily in highly vascularized tissues such as placenta, uterus, tumors, fetal liver, fetal spleen and also in the C7MCF7 cell line treated with estrogen.

10

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometritis, endometrial carcinoma, primary hepatocellular carcinoma, and spleen-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endometrium, liver and spleen, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endometrium, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 12 are useful for diagnosis and treatment of diseases of the endometrium (such as endometrial carcinoma, endometriosis, and endometritis), liver diseases (such as primary hepatocellular carcinoma), and spleen-related diseases.

20

SEQ ID NO: 145 as residues: Ala-29 to Leu-35, Leu-50 to Ser-57, Glu-96 to Glu-105, Asp-140 to Asp-148, and Asn-191 to Ser-197.

30

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 13

Gene NO: 13 is expressed primarily in B cell lymphoma and to a lesser extent in other tissues.

35

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B cell lymphoma; hematopoietic disorders; immune dysfunction.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may 5 routinely be detected in certain tissues and cell types (e.g., bone marrow and B-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the 10 disorder.

Enhanced expression of this gene product in B cell lymphoma indicates that it may play a role in the proliferation of hematopoietic cells. It is also believed to be involved in the survival and/or differentiation of various hematopoietic lineages. Expression in lymphoma also indicates that it may be involved in other cancers and 15 abnormal cellular proliferation. The tissue distribution, therefore, indicates that polypeptides and polynucleotides corresponding to Gene NO: 13 are useful for the diagnosis and/or therapeutic treatment of hematopoietic disorders, particularly B cell lymphoma. Furthermore, since overexpression of this gene is associated with the development of B cell lymphoma, antagonists of this protein are useful to interfere with 20 the progression of the disease. This protein is useful in assays for identifying such antagonists. Assays for identifying antagonists are known in the art and are described briefly elsewhere herein. Preferred antagonists include antibodies and antisense nucleic acid molecules. Preferred are antagonists which inhibit B-cell proliferation.

## 25 FEATURES OF PROTEIN ENCODED BY GENE NO: 14

The translation product of Gene NO: 14 shares sequence homology with very low density lipoprotein receptor which is thought to be important in transport of lipoproteins. Owing to the sequence similarity the translation product of this gene is believed to share certain biological activities with VLDL receptors. Assaying such 30 activity may be achieved by assays known in the art and set forth elsewhere herein.

This gene is expressed primarily in human synovium, umbilical vein endothelial cells, CD34+ cells, Jurkat cells, and HL60 cells, and to a lesser extent in thymus, meningioma, hypothalamus, adult testis, and fetal liver and spleen.

Therefore, polynucleotides or polypeptides of the invention are useful as 35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, atherosclerosis, ataxia malabsortion, vascular damage, hyperlipidemia,

and other cardiovascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and hematological systems,

- 5 expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endothelium, thymus meningoia, hypothalmus, testes, liver, and spleen and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,  
10 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in the vascular endothelial cells and homology to VLDL receptors indicates that polypeptides and polynucleotides corresponding to Gene NO: 14 are useful for diagnosis and treatment of atherosclerosis, ataxia malabsortion, and  
15 hyperlipidemia. These and other factors often result in other cardiovascular diseases. Additionally, the presence of the gene product in cells of blood lineages indicates that it may be useful in hematopoietic regulation and hemostasis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 147 as residues: Pro-39 to Ser-52, Trp-71 to Thr-76, and Pro-94 to His-100.

20

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 15

The translation product of Gene NO: 15 shares sequence homology with kallikrein which is thought to be important in blood pressure and renal secretion. Furthermore, this gene has now been characterized as a novel hepatitis B virus X  
25 binding protein that inhibits viral replication. See, for example, J. Virol. 72 (3), 1737-1743 (1998).

This gene is expressed primarily in kidney, placenta, lung, aorta and other endothelial cells, caudate nucleus and to a lesser extent in melanocytes, liver, adipose tissue.

30 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renovascular hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy. Similarly, polypeptides and antibodies directed to these  
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renovascular or respiratory vascular systems, expression of this gene

- at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., kidney, placenta, lung, endothelial cells, melanocytes, liver, and adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from 5 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to kallikrein indicates that polypeptides and polynucleotides corresponding to Gene NO: 15 are useful for treating renovascular hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy and 10 hydronephrosis. The protein expression in the organs like kidney, lung and vascular endothelial cells indicates the gene involvement in hemodynamic regulatory functions. The translation product of this gene is also useful in the treatment of viral infection, particularly liver infection, and particularly hepatitis B virus(es).

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 148 as residues: Leu-9 to Asn-15 and Thr-56 to Asp-61.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of Gene NO: 16 shares sequence homology with 20 secretory component protein, immunoglobulins and their receptors which are thought to be important in immunological functions. The amino acid sequence of secretory component protein can be accessed as accession no. pirlA02112, incorporated herein by reference.

Gene NO: 16 is expressed primarily in macrophages, monocytes and dendritic 25 cells and to a lesser extent in placenta and brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and tumors. Similarly, polypeptides and antibodies directed 30 to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cells (e.g., macrophages, monocytes, dendritic cells, placenta and brain, and cancerous 35 and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to immunoglobulins and secretory component protein indicates that polypeptides and polynucleotides corresponding to

5 Gene NO: 16 are useful for diagnosis and treatment of inflammation and bacterial infection, and other diseases where immunomodulation would be beneficial.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 149 as residues: Pro-37 to Cys-51, Gln-53 to Cys-60, Asn-99 to Gly-106, Gly-145 to Glu-151, and Ile-159 to Ser-164.

10

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of Gene NO: 17 is evolutionarily conserved and shares sequence homology with proteins from yeast and *C. elegans*. See, for example, Genbank accession no.gil746540. As is known in the art, strong sequence similarity to a secreted protein from *C. elegans* is predictive of cellular location of human proteins.

15 Gene NO: 17 is expressed primarily in colon carcinoma cell lines, messangial cells, many tumors like T cell lymphoma, osteoclastoma, Wilm's tumor, adrenal gland tumor, testes tumor, synovial sarcoma, and to a lesser extent in placenta, lung and brain.

20

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rapidly growing/dividing cells such as cancerous tissue, including, colon carcinoma, lymphomas, and sarcomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal, hematological and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, lung, brain, colon, messangial cells, 25 adrenal gland, T-cells, testes, and lymph tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30

35 The tissue distribution in colon cancer and many other tumors indicates that the polynucleotides and polypeptides of Gene NO: 17 are useful for cancer diagnosis and therapeutic targeting. The extracellular nature may contribute to solid tumor

immunosuppression, angiogenesis and cell growth stimulation. The tissue distribution of this gene in cells of the immune system indicates that polypeptides and polynucleotides corresponding to Gene NO: 17 are useful for treatment, prophylaxis and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, 5 allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. Its expression predominantly in hematopoietic cells also indicates that the gene could be important for the treatment and/or detection of hematopoietic disorders such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The protein can also 10 be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic 15 disorders including immunodeficiency diseases, leukemia, and septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 150 as residues: Val-131 to Asn-136.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 18

The translation product of Gene NO: 18 shares sequence homology with 20 immunoglobulin, which is thought to be important in immunoreactions. Gene NO: 18 is expressed primarily in macrophage. Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 25 biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or 30 lower levels may routinely be detected in certain tissues and cell types (e.g., macrophage and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the 35 disorder.

The tissue distribution in macrophages and the weak homology to immunoglobulin indicates that polypeptides and polynucleotides corresponding to Gene

NO: 18 are useful for diagnosing and treating immune response disorders, including inflammation, antigen presentation and immunosurveillance.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 19**

5       The translation product of Gene NO: 19 shares sequence homology with proline rich proteins which are thought to be important in protein-protein interaction.

This gene has a wide range of tissue distribution, but is expressed primarily in normal prostate, synovial fibroblasts, brain amygdala depression, fetal bone and fetal cochlea, and to a lesser extent in adult retina, umbilical vein endothelial cells, atrophic 10     endometrium, osteoclastoma, melanocytes, pancreatic carcinoma and smooth muscle.

10     Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer metastasis, wound healing, tissue repair. Similarly, polypeptides 15     and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal, connective tissues, reproductive and central nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, 20     prostate, fibroblasts, bone, cochlea, retina, endothelial cells, endometrium, pancreas and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., 25     the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to proline-rich proteins indicates that the protein is a extracellular matrix protein or an ingredient of bodily fluid. Polypeptides and polynucleotides corresponding to Gene NO: 19 are useful for cancer metastasis intervention, tissue culture additive, bone modeling, wound healing and tissue repair.

30

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 20**

Gene NO: 20 is expressed primarily in prostate cancer, leukocytes, meningima, adult liver, pancreas, brain, and to a lesser extent in lung.

35     Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancers. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., prostate,

5      leukocytes, meninges, liver, brain, pancreas and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10     Prostate cancer cell lines are known to be responsive to estrogen and androgen. The protein expression of Gene NO: 20 appears to be influenced by both estrogen and androgen levels. The prostate cancer tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 20 are useful in the intervention and detection of prostate hyperplasia and prostate cancer.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 21**

The translation product of Gene NO: 21 is identical to the human wnt-7a gene. Wnt-7a is a secreted signaling molecule, thought to be important in signaling and the regulation of cell fate and pattern formation during embryogenesis. Specifically, knock 20    out studies in mice have demonstrated that wnt7a plays a critical role in the development of the dorsal-ventral patterning in the developing limb, and to a lesser extent plays a role in the development of anterior-posterior patterning. Overexpression of wnt7a can induce transformation of cultured mammary cells, suggesting that it is an oncogene.

Expression of Gene NO: 21 has only been observed in testes.

25     Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, testicular cancer; abnormal limb development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological 30    probes for differential identification of the testes or developing embryo. For a number of disorders of the above tissues or cells, particularly of the developing embryo, expression of this gene at significantly higher or lower levels may routinely be detected in the developing embryo or amniotic fluid taken from a pregnant individual and compared relative to the standard gene expression level, i.e., the expression level in 35    healthy tissue or bodily fluid from an individual not having the disorder. Also, expression of this gene at significantly higher or lower levels may routinely be detected in the testes of patient suffering from testicular cancer and compared relative to the

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to mouse wnt7a indicates that polypeptides and polynucleotides corresponding to Gene NO: 21 are useful to restore abnormal limb development in an affected individual. Furthermore, its oncogenic potential and tissue distribution indicates that it could serve as a diagnostic for testicular cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 154 as residues: Gly-22 to Arg-28.

10

## FEATURES OF PROTEIN ENCODED BY GENE NO: 22

Gene NO: 22 is expressed primarily in fetal liver/spleen, breast, testes and placenta and to a lesser extent in brain, and a series of cancer tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, brain diseases, male infertility, and disposition to pregnant miscarriages. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoietic system, and sexual organs, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, testes, placenta, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 22 are useful as a marker for non-differentiated, dividing cells and hence could serve as an oncogenic marker. Its high expression in fetal liver, suggests an involvement in hematopoiesis and/or the immune system. Hence it is useful as a factor to enhance an individuals immune system, e.g., in individuals with immune disorders. It is also thought to affect the survival, proliferation, and differentiation of a number of hematopoietic cell lineages, including hematopoietic stem cells. Its disruption, e.g., mutation or altered expression, may also be a marker of immune disorder. Its expression in the testes, suggests it may be important in controlling male fertility. Expression of this gene in breast further reflects a

role in immune function and immune surveillance (breast lymph node). This gene is believed to be useful as a marker for breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 155 as residues: Gln-57 to Lys-70 and Ala-91 to Pro-100.

5

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 23**

Gene NO: 23 is expressed primarily in bone marrow and brain (whole and fetal).

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and hematopoietic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 23 are useful in the diagnosis and treatment of disorders related to the central nervous system (e.g. neuro-degenerative conditions, trauma, and behavior abnormalities) and hematopoiesis. In addition, the expression in fetal brain indicates a role for this gene product in diagnosis of predisposition to developmental defects of the brain.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 156 as residues: Thr-23 to Tyr-29.

30

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 24**

Gene NO: 24 is expressed primarily in smooth muscle, placenta, prostate, and osteoblasts.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular pathologies. Similarly, polypeptides and antibodies

directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, reproductive and skeletal systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, smooth muscle, prostate, and osteoblasts, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 24 are useful for detection and treatment of neoplasias and developmental abnormalities associated with these tissues.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 157 as residues: Asn-21 to Thr-26.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of Gene NO: 25 shares sequence homology with Pregnancy Associated Mouse Protein (PAMP)-1. (See, FEBS Lett 1993 May 20 17;322(3):219-222). Based on the sequence similarity the translation product of this gene is expected to share certain biological activities with PAMP-1.

Gene NO: 25 is expressed primarily in 12-week-old human embryos and prostate.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate disorders (cancer). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 25 are useful for the diagnosis and treatment of prostate disorders (such as cancer) and developmental abnormalities and fetal deficiencies. The homology to PAMP-1 indicates that this gene and gene product are useful in detecting 5 pregnancy.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 158 as residues: Pro-23 to Glu-28 and Ser-44 to Gly-55.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 26

10 Gene NO: 26 is expressed primarily in testes and to a lesser extent in epididymis.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 15 not limited to, reproductive and endocrine disorders, as well as testicular cancer.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive and endocrine systems, expression of this gene at significantly higher or 20 lower levels may routinely be detected in certain tissues (e.g., testes, and epididymis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the 25 disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 26 are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility 30 and/or impotence. This gene product is also useful in assays designed to identify binding agents as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, 35 this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 159 as residues: Pro-24 to Gly-33 and Arg-70 to Gly-76.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 27**

- 5       The translation product of Gene NO: 27 shares sequence homology with salivary protein precursors which are thought to be important in immune response and production of secreted proteins.
- Gene NO: 27 is expressed primarily in salivary gland tissue.
- Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, diseases of the salivary gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, digestive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., salivary gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution and homology to salivary secreted protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 27 are useful for treatment of immune disorders and diagnostic uses related to secretion of protein in disease states. For example, the gene product can be used as an anti-microbial agent, an ingredient for oral or dental hygiene, treatment of xerostomia, sialorrhea, intervention for inflammation including parotitis, and an indication for tumors in the salivary gland (adenomas, carcinomas).

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 160 as residues: Asp-21 to Gly-28, Asp-30 to Glu-43, Glu-49 to Glu-62, and Thr-75 to Pro-83.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 28**

- Gene NO: 28 is expressed primarily in human fetal heart tissue and to a lesser extent in olfactory tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, olfactory and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of 5 disorders of the above tissues or cells, particularly of the immune, olfactory and vascular systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., olfactory tissue, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a 10 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 28 are useful for diagnosis and treatment of immune, olfactory and vascular disorders.

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 161 as residues: Cys-33 to Gly-44, Arg-71 to Arg-78, Ser-130 to Gly-142, Lys-150 to Gly-157, and Thr-159 to Asp-177.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 29

20 Gene NO: 29 is expressed primarily in brain and to a lesser degree in activated macrophages, endothelial and smooth muscle cells, and some bone cancers.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of brain and endothelial present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited 25 to, neurodegeneration, inflammation and other immune disorders, fibrotic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification brain, smooth muscle, and endothelium. For a number of disorders of the above tissues or cells, particularly of the brain and endothelium, expression of this gene at significantly higher or lower 30 levels may routinely be detected in certain tissues or cell types (e.g., brain, endothelial cells, macrophages, smooth muscle, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily 35 fluid from an individual not having the disorder.

Tissue distribution suggests polypeptides and polynucleotides corresponding to Gene NO: 29 are useful in study and treatment of neurodegenerative and immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:

- 5 162 as residues: Asn-18 to Glu-20, Ser-33 to Gln-48, Cys-55 to Ser-56, Pro-67 to Cys-69.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 30**

- Gene NO: 30 is expressed primarily in early stage human brain and to a lesser 10 extent in cord blood, heart, and some tumors.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of developing CNS tissue present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular and neurodegenerative disorders. Similarly, polypeptides and 15 antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and heart, and cancerous and wounded tissues) or bodily 20 fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that that polypeptides and polynucleotides 25 corresponding to Gene NO: 30 are useful for the treatment of cancer and of neurodegenerative and cognitive disorders.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 31**

- Gene NO: 31 is expressed primarily in brain and thymus and to a lesser extent 30 in several other organs and tissues including the hematopoietic system, liver skin and bone

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are 35 not limited to, CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g.,

5      hematopoietic cells, brain, thymus, liver, bone, and epidermis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10     The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 31 are useful for treatment and diagnosis of CNS disorders, hematopoietic system disorders, disorders of the endocrine system, and of bone and skin.

15     Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 164 as residues: Thr-35 to Arg-40, Pro-55 to His-75, Pro-93 to Ala-98, Ala-111 to Pro-119, and Pro-132 to Glu-138.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 32

Gene NO: 32 is expressed primarily in organs and tissue of the nervous system  
20     and to a lesser extent in various developing tissues and organs.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system and disorders of developing and growing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly disorders of the CNS, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., tissue of the nervous system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35     The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 32 are useful for diagnosis and treatment of disorders of the central nervous system, general neurological diseases and neoplasias.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 165 as residues: Ser-33 to Lys-41 and Glu-86 to Glu-91.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 33

5 Residues 141-156 in the translation product for Gene NO: 33 as shown in the sequence listing matches phosphopantetheine binding site motifs. Phosphopantetheine (or pantetheine 4' phosphate) is the prosthetic group of acyl carrier proteins (ACP) in some multienzyme complexes where it serves as a 'swinging arm' for the attachment of activated fatty acid and amino-acid groups. Phosphopantetheine is attached to a serine  
10 residue in these proteins. ACP proteins or domains have been found in various enzyme systems which are listed below. Fatty acid synthetase (FAS), which catalyzes the formation of long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH. Bacterial and plant chloroplast FAS are composed of eight separate subunits which correspond to the different enzymatic activities; ACP is one of these polypeptides.  
15 Fungal FAS consists of two multifunctional proteins, FAS1 and FAS2; the ACP domain is located in the N-terminal section of FAS2. Vertebrate FAS consists of a single multifunctional enzyme; the ACP domain is located between the beta-ketoacyl reductase domain and the C-terminal thioesterase domain. Based on the presence of a phosphopantetheine binding site in the translation product of this gene, it is believed to  
20 share activities fatty acid synthetase polypeptides. Such activities may be assayed by methods known in the art.

This gene is expressed primarily in developing and rapidly growing tissues like placenta fetal heart and endometrial tumor and to a lesser extent in B and T cell lymphoma tissues

25 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and disorders of developing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing  
30 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic tissues and developing organs and tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., embryonic tissue, endometrium, B-cells, and T-cells, and cancerous and wounded  
35 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 33 are useful for treatment and diagnosis of cancer in the hematopoietic system developing organs and tissues. It may also be useful for induction of cell growth in disorders of the hematopoietic system and other tissue and organs. The homology to fatty acid synthetases indicates that this gene product is useful in the diagnosis and treatment of lipid metabolism disorders such as hyperlipidemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 10 166 as residues: Arg-27 to Glu-34.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 34

Gene NO: 34 is expressed primarily in breast and testes tissues and to a lesser extent in hematopoietic tissues including tonsils, T cells and monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the reproductive organs and systems, including cancer, autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive organs and hematopoietic tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells, T-cells and monocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Nucleic acids comprising sequence of this gene are also useful as chromosome markers since this gene maps to Chr.15, D15S118-D15S123.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 34 are useful for treatment of diseases of the reproductive organs and hematopoietic system including cancer, autoimmune diseases and inflammatory diseases .

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 167 as residues: Phe-81 to Lys-86.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 35**

The translation product of Gene NO: 35 shares sequence similarity with the mouse cytokine-inducible inhibitor of signaling. See, e.g., Nature 1997 Jun

- 5      26;387(6636):917-921. Cytokines are secreted proteins that regulate important cellular responses such as proliferation and differentiation. Key events in cytokine signal transduction are well defined: cytokines induce receptor aggregation, leading to activation of members of the JAK family of cytoplasmic tyrosine kinases. In turn, members of the STAT family of transcription factors are phosphorylated, dimerize and
- 10     increase the transcription of genes with STAT recognition sites in their promoters. Less is known of how cytokine signal transduction is switched off. Expression of the mouse SOCS-1 protein inhibited both interleukin-6- induced receptor phosphorylation and STAT activation. We have also cloned two relatives of SOCS-1, named SOCS-2 and SOCS-3, which together with the previously described CIS form a new family of
- 15     proteins. Transcription of all four SOCS genes is increased rapidly in response to interleukin-6, in vitro and in vivo, suggesting they may act in a classic negative feedback loop to regulate cytokine signal transduction. The translation product of this gene is believed to have similar biological activities as this family of mouse genes. The biological activity of the translation product of this gene may be assayed by methods
- 20     shown in Nature 1997 Jun 26;387(6636): 917-921, which is incorporated herein by reference in its entirety.

Gene NO: 35 is expressed primarily in tissues of hematopoietic origin including activated monocytes, neutrophils, activated T-cells and to a lesser extent in breast, adipose tissue and dendritic cells.

- 25     Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the hematopoietic system including cancer autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to
- 30     these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells and cancerous and wounded tissues) or bodily fluids
- 35     (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to cytokine inducible inhibitor of signaling indicates that polypeptides and polynucleotides corresponding to Gene NO: 35 are

5 useful for diagnosis and treatment of diseases of the hematopoietic system including autoimmune diseases, inflammatory diseases, infectious diseases and neoplasia. For example, administration of, or upregulation of this gene could be used to decrease the response of immune-system to lymphokines and cytokines.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 10 168 as residues: Arg-23 to His-30, Ala-35 to Gly-42.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 36**

Gene NO: 36 is expressed primarily in infant brain and to a lesser extent in osteoclastoma, placenta, and a wide variety of other tissues.

15 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential 20 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., osteoclastoma, placenta, and tissue of the central nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial 25 fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 36 are useful for diagnosis and treatment of neurologic 30 disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 169 as residues: Gln-31 to Ser-37, Ile-49 to Gly-54, Tyr-57 to Asp-67, Gln-141 to Pro-151, and Val-207 to Thr-219.

#### **35 FEATURES OF PROTEIN ENCODED BY GENE NO: 37**

Gene NO: 37 is expressed primarily in osteoclastoma stromal cells, dendritic cells, liver, and placenta.

- Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, wound, pathological conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., stromal cells, dendritic cells, liver, and placenta and, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 37 are useful for fundamental role in basic growth and development of human.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 170 as residues: Leu-32 to Thr-37 and Arg-48 to Pro-55.

## 20 FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of Gene NO: 38 shares sequence homology with a yeast protein, Lpe10p, which may be involved in mRNA processing. (See Accession Nos. 2104457 and 1079682.) It is likely that an upstream signal sequence exists, other than the predicted sequence described in Table 1. Preferred polypeptide fragments comprise the open reading frame upstream from the predicted signal sequence, as well as polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in skin, and to a lesser extent in embryonic tissues, and fetal liver.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, liver, and embryonic tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum,

plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5       The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 38 are useful for diagnosis and treatment of defects of the skin.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 39**

10      Gene NO: 39 is expressed primarily in Amygdala, activated monocytes, testis, and fetal liver. Moreover, this gene is mapped to chromosome 4. Thus, polynucleotides of the present invention can be used in linkage analysis as markers for chromosome 4.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the brain, immune system and testis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, immune system and testis, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., Amygdala, monocytes, testes, and liver and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 39 are useful for detecting defects of the brain, immune system and testis because of its abundance in these tissues.

30

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 40**

The translation product of Gene NO: 40 shares sequence homology with lymphoma 3-encoded protein (bcl-3) which is thought to contribute to leukemogenesis when abnormally expressed.

35      This gene is expressed primarily in Human Neutrophils, and to a lesser extent in Human Osteoclastoma Stromal Cells (unamplified), Hepatocellular Tumor, and Human Neutrophils, (Activated).

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., neutrophils, osteoclastoma, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to lymphoma 3-encoded protein (bcl-3) indicates that polypeptides and polynucleotides corresponding to Gene NO: 40 are useful for treatment of lymphoma and related cancers.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 41

Gene NO: 41 is expressed primarily in ovary tumor, and to a lesser extent in endometrial stromal cells and fetal brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, ovarian or endometrial cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system and the developing central nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., ovary, endometrium and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 41 are useful for development of factors involved in ovarian or endometrial and general reproductive organ disorders.

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 5 174 as residues: Glu-22 to Trp-31, Asn-84 to Asp-90, and Ser-144 to Asp-151.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 42**

The translation product of Gene 42 has sequence identity with a gene designated PTHrP(B). The PTHrP(B) polypeptide inhibits parathyroid hormone related peptide 10 (PTHrP) activity.

This gene is expressed primarily in adult testis, and to a lesser extent in pituitary.

Therefore polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a 15 biological sample and for diagnosis of male reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may routinely be 20 detected in certain tissues (e.g., testes, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Furthermore, based in part on 25 sequence identity with PTHrP(B), nucleic acids and polypeptides of the present invention may be used to diagnose or treat such conditions as hypercalcemia, osteoporosis, and disorders related to calcium metabolism.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 42 are useful for treatment of male reproductive disorders, 30 hypercalcemia, osteoporosis, and other disorders related to calcium metabolism.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 175 as residues: Tyr-81 to Met-86, Gly-103 to Ser-108, Glu-127 to Pro-128, Pro-175 to Ser-180, Glu-196 to Lys-203, Pro-235 to Ser-241, and Ala-249 to Ser-264.

#### **35 FEATURES OF PROTEIN ENCODED BY GENE NO: 43**

The translation product of Gene NO: 43 shares sequence homology with brevican, which is thought to be important as a proteoglycan core protein of the

aggregcan/versican family. The translation product of this gene may also contain a hyaluronan (HA)-binding region domain in frame with, but downstream of, the predicted open reading frame (Barta, et al., Biochem. J. 292:947-949 (1993)). The HA-binding domain, also termed the link domain, is found in proteins of vertebrates  
5 that are involved in the assembly of extracellular matrix, cell adhesion, and migration. It is about 100 amino acids in length. The structure has been shown to consist of two alpha helices and two antiparallel beta sheets arranged around a large hydrophobic core similar to that of C-type lectin. This domain typically contains four conserved cysteines involved in two disulfide bonds.

10 This gene is expressed primarily in early stage human brain and to a lesser extent in frontal cortex and epileptic tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of disorders associated with, or observed during,  
15 neuronal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of neuronal and associated tissues and cell types. For a number of disorders of the above tissues or cells, particularly for those of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g.,  
20 brain and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution and homology to brevican indicates that polypeptides and polynucleotides corresponding to Gene NO: 43 are useful for neuronal regulation and signaling. The uses include directing or inhibiting axonal growth for the treatment of neuro-fibromatosis and in detection of glioses.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:  
30 176 as residues: Asp-28 to Arg-33 and Arg-126 to Arg-131.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 44

Gene NO: 44 is the human homolog of Notch-2 (Accession No. 477495) and mouse EGF repeat transmembrane protein (Accession No. 1336628), both genes are  
35 important in differentiation and development of an organism. The EGF repeat transmembrane protein is regulated by insulin like growth factor Type I receptor. These proteins are involved in cell-cell signaling and cell fate determination. Based on

homology, it is likely that this gene products also involved in cell differentiation and development. Although the predicted signal sequence is indicated in Table 1, it is likely that a second signal sequence is located further upstream. Moreover, further translated coding regions are likely found downstream from the disclosed sequence, which can  
5 easily be obtained using standard molecular biology techniques. A frameshift occurs somewhere around nucleotide 714, causing a frame shift in amino acid sequence from frame +2 to frame +3. However, using the homology of Notch-2 and EGF repeat transmembrane protein, the complete open reading frame can be elucidated. Preferred polynucleotide fragments comprise nucleotides 146-715, 281-715, and 714-965. Other  
10 preferred polypeptide fragments comprise the following EGF-like motifs:  
CRCASGFTGEDC (SEQ ID NO:260), CTCQVGFTGKEC (SEQ ID NO:261),  
CLNLPGSYQCQC (SEQ ID NO:262), CKCLTGFTGQKC (SEQ ID NO:263), and  
CQCLQGFTGQYC (SEQ ID NO:264).

Gene NO: 44 is expressed primarily in placenta and to a lesser extent in stromal  
15 and immune cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hemophilia and other blood disorders, central nervous system disorders,  
20 muscle disorders, and any other disorder resulting from abnormal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and vascular systems, expression of this gene at significantly  
25 higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, stromal and immune cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an  
30 individual not having the disorder.

The tissue distribution and homology to Notch-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 44 are useful for diagnosing and treating disorders relating to abnormal regulation of cell fate, induction, and differentiation of cells (e.g., cancer), epidermal growth factors, axonal pathfinding, and hematopoiesis.

35 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 177 as residues: Gln-27 to Tyr-32, His-45 to Glu-55, Tyr-61 to Gly-77, Glu-99 to Ser-106, Ser-125 to Cys-131, and Thr-138 to Trp-144.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 45**

The translation product of this gene shares sequence homology with Laminin A which is thought to be important in the binding of epithelial cells to basement membrane 5 and is associated with tumor invasion. Moreover, the translated protein is homologous to the *Drosophila* LAMA gene (Accession No. 1314864), a gene expressed in the first optic ganglion of *Drosophila*. Thus, it is likely that the gene product from this gene is involved in the development of the eye. Nucleotide fragments comprising nucleotides 822-1223, 212-475, 510-731, and 1677-1754 are preferred. Also preferred are the 10 polypeptide fragments encoded by these polynucleotide fragments. It is likely that a frame shift occurs somewhere between nucleotides 475 to 510, shifting the open reading frame from +2 to +3. However, the open reading frame can be clarified using known molecular biology techniques.

This gene is expressed primarily in human testes tumor and to a lesser extent in 15 placenta and activated monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, invasive cancers or tumors of the epithelium, as well as disorders relating 20 to eye development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of neoplastic conditions, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., testes, 25 placenta, and monocytes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The tissue distribution and homology to Laminin A indicates that polypeptides and polynucleotides corresponding to Gene NO: 45 are useful for study and diagnosis of malignant or benign tumors, fibrotic disorders, and eye disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 178 as residues: Met-1 to Gly-8, Glu-32 to Ala-37, Met-113 to Asn-119, and Glu-139 35 to Gln-153.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 46**

The translation product of Gene NO: 46 is novel and shares sequence homology with the product of the *Drosophila* tissue polarity gene frizzled. In vertebrates, it appears that there is a family of proteins that represent frizzled gene homologs. (See, e.g., Accession Nos. 1946343 and AFO17989.) The *Drosophila* frizzled protein is thought to transmit polarity signals across the plasma membrane of epidermal cells. The structure of frizzled proteins suggest that they may function as a G-protein-coupled receptor. The frizzled proteins are thought to represent receptors for Wnt gene products - secreted proteins that control tissue differentiation and the development of embryonic and adult structures. Inappropriate expression of Wnts has also been demonstrated to contribute to tumor formation. Moreover, mammalian secreted frizzled related proteins are thought to regulate apoptosis. (See Accession No. AFO17989.) The human homolog has also been recently cloned by other groups. (See Accession No. H2415415.) Thus, the protein encoded by this gene plays a role in mediating tissue differentiation, proliferation, tumorigenesis and apoptosis. Preferred polypeptide fragments lack the signal sequence as described in Table 1, as well as N-terminal and C-terminal deletions. Preferred polynucleotide fragments encode these polypeptide fragments.

Gene NO: 46 is expressed primarily in fetal tissues - particularly fetal lung - and adult cancers, most notably pancreas tumor and Hodgkin's lymphoma. Together, this distribution is consistent with expression in tissues undergoing active proliferation. The gene is also expressed to a lesser extent in other organs, including stomach, prostate, and thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer (particularly pancreatic cancer and/or Hodgkin's lymphoma), as well as other forms of aberrant cell proliferation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and hyperproliferative disorders, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., fetal tissue, pancreas, and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to frizzled indicates that polypeptides and polynucleotides corresponding to Gene NO: 46 are useful for influencing cell proliferation, differentiation, and apoptosis. The full-length protein or a truncated domain could potentially bind to and regulate the function of specific factors, such as Wnt proteins or other apoptotic genes, and thereby inhibit uncontrolled cellular proliferation. Expression of this protein within a cancer - such as via gene therapy or systemic administration - could effect a switch from proliferation to differentiation, thereby arresting the progression of the cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 179 as residues: Pro-31 to Arg-37.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of Gene NO: 47 shares sequence homology with members of the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes. These ribonuclease proteins are found predominantly in fungi, plants, and bacteria and have been implicated in a number of functions, including phosphate-starvation response, self-incompatibility, and responses to wounding. A second group has recently cloned this same gene, calling it a ribonuclease 6 precursor. (See Accession No. 2209029.) This group also mapped the gene to chromosome 6, thus, the polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 6.

Gene NO: 47 is expressed primarily in hematopoietic cells and tissues, including macrophages, eosinophils, CD34 positive cells, T-cells, and spleen. It is also expressed to a lesser extent in brain and spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a hematopoietic origin, graft rejection, wounding, inflammation, and allergy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes indicates that polypeptides and polynucleotides

- 5 corresponding to Gene NO: 47 are useful as a cytotoxin that could be directed against specific cell types (e.g. cancer cells; HIV- infected cells), and that would be well tolerated by the human immune system.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 180 as residues: Ala-24 to Asp-30, Ile-51 to Tyr-61, Pro-69 to Ser-78, Pro-105 to Phe-110, Asn-129 to Phe-135, Pro-187 to Glu-192, Lys-205 to Gln-224, and Pro-250 to His-256.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of Gene NO: 48 shares sequence homology with dolichyl-phosphate glucosyltransferase, a transmembrane-bound enzyme of the endoplasmic reticulum which is thought to be important in N-linked glycosylation, by catalyzing the transfer of glucose from UDP-glucose to dolichyl phosphate. (See Accession No. 535141.) Based on homology, it is likely that this gene product also play a role similar in humans. Preferred polynucleotide fragments comprise nucleotides 132-959. Also preferred are the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 48 is expressed primarily in endothelial cells and to a lesser extent in hematopoietic cells and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects in proper N-linked glycosylation of proteins, such as Wiskott-Aldrich syndrome; tumors of an endothelial cell origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and hematopoietic systems, as well as brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., endothelial cells, hematopoietic cells, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to dolichyl-phosphate glucosyltransferase indicates that polypeptides and polynucleotides corresponding to Gene NO: 48 are 5 useful in diagnosing and treating defects in N-linked glycosylation pathways that contribute to disease conditions and/or pathologies.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 181 as residues: Lys-50 to Thr-55, Ser-73 to Arg-79, Glu-92 to Pro-99, Asp-110 to Ser-117, Gln-125 to Lys-131, Gly-179 to Asn-188, Ile-231 to Cys-236, and Glu-318 10 to Asn-324.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 49**

Gene NO: 49 is expressed primarily in brain, most notably in the hypothalamus and amygdala. This gene is also mapped to chromosome X, and therefore, can be used 15 in linkage analysis as a marker for chromosome X.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a brain origin; neurodegenerative disorders, and sex-linked 20 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and cancerous and wounded tissues) or bodily 25 fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides 30 corresponding to Gene NO: 49 are useful for the diagnosis of tumors of a brain origin, and the treatment of neurodegenerative disorders, such as Parkinson's disease, and sex-linked disorders.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 50**

The translation product Gene NO: 50 shares sequence homology with canine phospholemman, a major plasma membrane substrate for cAMP-dependent protein 35 kinases A and C. (See Accession No. M63934; see also Accession No. A40533.) In

fact, a group also recently cloned the human phospholemman gene, and mapped this gene to chromosome 19. (See Accession No.1916010.) Phospholemman is a type I integral membrane protein that gets phosphorylated in response to specific extracellular stimuli such as insulin and adrenalin. Phospholemman forms ion channels in the cell membrane and appears to regulate taurine transport, suggesting an involvement in cell volume regulation. It has been proposed that phospholemman is a member of a superfamily of membrane proteins, characterized by single transmembrane domains, which function in transmembrane ion flux. They are capable of linking signal transduction to the regulation of such cellular processes as the control of cell volume.

Gene No 50 is expressed primarily in fetal liver and to a lesser extent in adult brain and kidney, as well as other organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, insulin and/or adrenalin defects; diabetes; aberrant ion channel signaling; defective taurine transport; and defects in cell volume regulation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and/or immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, brain, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to phospholemman indicates that polypeptides and polynucleotides corresponding to Gene NO: 50 are useful for treatment of disorders involving the transport of ions and small molecules, in particular taurine. It could also be beneficial for control of pathologies or diseases wherein aberrancies in the control of cell volume are a distinguishing feature, due to the predicted role for phospholemman in the normal control of cell volume. It also may play a role in disorders involving abnormal circulating levels of insulin and/or adrenalin - along with other active secreted molecules - as revealed by its phosphorylation upon stimulation with insulin or adrenalin.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 183 as residues: Ala-20 to Gln-34, Arg-58 to Thr-79, and Leu-87 to Arg-92.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 52**

Gene NO: 52 is expressed primarily in metastatic melanoma and to a lesser extent in infant brain.

- 5 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and cancer metastasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a
- 15 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 52 are useful for diagnosis and treatment of melanoma.

**20 FEATURES OF PROTEIN ENCODED BY GENE NO: 53**

The translation product of Gene NO: 53 shares sequence homology with mucin which is thought to be important cell surface molecule. It also exhibits sequence identity with a calcium channel blocker of *Agelenopsis aperta*. In particular, with those calcium channel blockers which affect neuronal and muscle cells.

- 25 Gene NO: 53 is expressed primarily in prostate, endothelial cells, smooth muscle and fetal tissues and to a lesser extent in T cells and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancer, immune disorders, angina, hypertension, 30 cardiomyopathies, supraventricular arrhythmia, oesophageal achalasia, premature labour, and Raynaud's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or 35 lower levels may routinely be detected in certain tissues or cell types (e.g., prostate, and tissue and cells of the immune system, and cancerous and wounded tissues) or

bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

5       The tissue distribution and homology to mucin indicates that polypeptides and polynucleotides corresponding to Gene NO: 53 are useful as a surface antigen for diagnosis of diseases such as prostate cancer and as tumor vaccine.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 54

10      Gene NO: 54 encodes a polypeptide which exhibits sequence identity with the rab receptor and VAMP-2 receptor proteins. (Martincic, et al., J. Biol. Chem. 272 (1997).)

Gene NO: 54 is expressed primarily in placenta, fetal liver, osteoclastoma and smooth muscle and to a lesser extent in T cell, fetal lung and colon cancer.

15      Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, osteoporosis and immuno-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing 20 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoiesis system and bone system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, liver, osteoclastoma, smooth muscle, T-cells, and lung, and colon, and 25 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30      The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 54 are useful for treating cancer, osteoporosis and immuno-disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 187 as residues: Pro-16 to Phe-21, Pro-24 to Arg-35, Arg-92 to Pro-98, Asn-143 to 35 Lys-151, and Leu-169 to Ile-176.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 55**

Gene NO: 55 encodes a protein having sequence identity to the rat galanin receptor GALR2.

Gene NO: 55 is expressed primarily in ovarian cancer.

- 5 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of ovarian cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders 10 of the above tissues or cells, particularly of the immune system and reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., ovary, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample 15 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. GALR2 antagonists can be used to treat obesity, bulimia, or Alzheimer's disease, while GALR2 agonists can be used to treat anorexia or pain, or to decrease nociception (claimed). Agonists and antagonists can also be used to 20 treat numerous other disorders, including cognitive disorders, sensory disorders, motion sickness, convulsion/epilepsy, hypertension, diabetes, glaucoma, reproductive disorders, gastric and intestinal ulcers, inflammation, immune disorders, and anxiety.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 55 are useful for diagnosis and treatment of ovarian cancer.

25

**FEATURES OF PROTEIN ENCODED BY GENE NO: 56**

- As indicated in Table 1, the predicted signal sequence of Gene NO: 56 relates to an open reading frame that is homologous to the mouse major histocompatibility locus class III. (See Accession No. 2564953.) Any frame shift mutations that alter the correct 30 open reading frame can easily be clarified using known molecular biology techniques. Moreover, in the opposite orientation, a second translated product is disclosed. This second translation product of this contig is identical in sequence to intracellular protein lysophosphatidic acid acyltransferase. The nucleotide and amino acid sequences of this translated product have since been published by Stamps and colleagues (Biochem. J. 326 (Pt 2), 455-461 (1997)), West and coworkers (DNA Cell Biol. 6, 691-701 (1997)), Rowan (GenBank Accession No. U89336), and Soyombo and Hofmann (GenBank Accession No. AF020544). This gene is thought to enhance cytokine

signaling response in cells. It is likely that a signal peptide is located upstream from this translated product. Preferred polypeptide fragments comprise the amino acid sequence: GLACWLAGVIFIDRKRTGDAISVMSEVAQTLTQDVXWVFPEGTRNHNGSML PFKRGAFHLAVQAQVPIVPMSSYQDFYCKKERRFTSGQCQVRVLPPVPTEGL

- 5 TPDVPALADRVRHSMLHCF(SEQ ID NO: 265);  
PSAKYFFKMAFYNGWILFLAVLAIPVCAVRGRNVENMKILRLMLLHIKYLYGI  
RVEVRGAHHFPPSQPYVVVSNHQSSL DLLGMMEVLPGRCPVIAKR (SEQ ID  
NO:266); TVFREISTD (SEQ ID NO:267); or LWAGSAGWPAG (SEQ ID NO: 268).  
Also provided are polynucleotide fragments encoding these polypeptide fragments.

10 Gene NO: 56 is expressed primarily in infant adrenal gland, hypothalamus, 7 week old embryonic tissue, fetal lung, osteoclastoma stromal cells, and to a lesser extent in a large number of additional tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of developmental disorders and osteoclastoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s) in which it is highly expressed. For a number of disorders of the above tissues or cells, particularly during development or of the nervous or bone systems, expression 20 of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., adrenal, embryonic tissue, lung, and osteoclastom al stromal cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Further, expression of this protein can be used to 25 alter the fatty acid composition of a given cell or membrane type.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 56 are useful for diagnosis and treatment of osteoclastoma and other bone and non-bone-related cancers, as well as for the diagnosis and treatment 30 of developmental disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 189 as residues: Gly-29 to Gly-36 and Tyr-49 to Tyr-58.

### 35 FEATURES OF PROTEIN ENCODED BY GENE NO: 57

The translation product of Gene NO: 57 shares sequence homology with longevity-assurance protein-1. (See Accession No. g1123105.) Preferred

polynucleotide fragments comprise nucleotides 6-125 and 118-432, as well as the polypeptides encoded by these polynucleotides. It is likely that a second signal sequence exists upstream from the predicted signal sequence in Table 1. Moreover, a frame shift likely occurs between nucleotides 118-125, which can be elucidated using standard molecular biology techniques.

Gene NO: 57 is expressed primarily in fetal liver, kidney, brain, thymus, and bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunological diseases and hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal liver, kidney, brain, thymus, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, kidney, brain, thymus, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to longevity-assurance protein suggest that Gene NO: 57 encodes a protein useful in increasing life span and in replacement therapy for those suffering from immune system disorders or hyperproliferative disorders caused by underexpression or overexpression of this gene.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 190 as residues: Val-29 to Arg-46 and Gly-50 to Gly-56.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 58**

Domains of the Gene NO: 58 product are homologous to porcine surfactant protein-A receptor. (See Accession No. B48516.) The bovine gene binds surfactant protein-A receptor, modulating the secretion of alveolar surfactant. Based on this homology, the gene product encoded by this gene will likely have activity similar to the porcine gene. Preferred polynucleotide fragments comprise nucleotides 887-1039, as well as the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 58 is expressed primarily in brain and to a lesser extent in endothelial cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the central nervous system including dementia, stroke, 5 neurological disorders, respiratory distress, and diseases affecting the endothelium including inflammatory diseases, restenosis, and vascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the placenta, liver, 10 endothelial cells, prostate, thymus, and lung, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene 15 expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology indicates that polypeptides and polynucleotides corresponding to Gene NO: 58 are useful for the diagnosis and /or treatment of diseases of the central nervous system, such as a factor that promote 20 neuronal survival or protection, in the treatment of inflammatory disorders of the endothelium, or in disorders of the lung. In addition this protein may inhibit or promote angiogenesis and therefore is useful in the treatment of vascular disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 191 as residues: His-66 to Pro-80, Gly-139 to Ser-146 and Ser-262 to Pro-267.

25

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The translation product of Gene NO: 59 is homologous to the rat hypertension-induced protein which is thought to be important in hypertension, and found expressed mainly in kidneys. (See Accession No. B61209.) Thus, it is likely that this gene 30 product is involved in hypertension in humans. Preferred polypeptide fragments comprise the short chain dehydrogenase/reductase motif SILGIISVPLSIGYCASKHALRGFFNGLR (SEQ ID NO:269), as well as polynucleotides encoding this polypeptide fragment. Also preferred are polynucleotide fragments of 337-639, as well as the polypeptide fragments encoded by this 35 polynucleotide fragment.

Gene NO: 59 is expressed primarily in liver, spleen, lung, brain, and prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular, immunological, and renal disorders. Similarly,

- 5 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, renal, and immune, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, lung, brain, and prostate, 10 and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 15 The tissue distribution and homology to hypertension-induced protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 59 are useful for treating hypertension.

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 192 as residues: Gln-40 to Glu-45, Glu-96 to Glu-102, Asn-256 to Thr-266, and Asp-20 308 to Asp-317.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 60

- Gene NO: 60 is expressed primarily in activated T-cell and jurkat cell and to a lesser extent in apoptic T-cell and CD34+ cell. It is likely that alternative open reading 25 frames provide the full length amino acid sequence, which can be verified using standard molecular biology techniques.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T lymphocyte related diseases or hematopoiesis. Similarly, polypeptides 30 and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain 35 tissues and cell types (e.g., T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 60 are useful for diagnosis or treatment of immune system disorders.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 61

The translation product of Gene NO: 61, a vacuolar proton-ATPase, shares sequence homology with a *Caenorhabditis elegans* protein which is thought to be important in development. This protein may be a human secretory homologue that may also influence embryo development. Ludwig, J., also recently cloned this gene from chromaffin granules. (See, Accession No. 2584788.) Although Table 1 indicates the predicted signal peptide sequence, the translated product of this gene may in fact start with the upstream methionine, beginning with the amino acid sequence MAYHGLTV (SEQ ID NO:270). Thus, polypeptides comprising this upstream sequence, as well as N-terminus deletions, are also contemplated in the present invention.

Gene NO: 61 is expressed primarily in human placenta, liver, and Hodgkin's Lymphoma and to a lesser extent in bone marrow. Modest levels of expression were also observed in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hyperproliferative disorders, defects in embryonic development, and diseases or disorders caused by defects in chromaffin granules. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., placenta, liver, lymph tissue, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* indicates that polypeptides and polynucleotides corresponding to Gene NO: 61 are useful for diagnostic or therapeutic modalities for hyperproliferative disorders, embryonic development disorders, and chromaffin granules disorders.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 62**

The translation product of Gene NO: 62 shares sequence homology with the murine LAG3 gene which is thought to be important in the mediation of natural killer cell (NK cell) activity as previously determined by experiments in mice containing null mutations of LAG3. The similarity of this gene to the CD4 receptor may imply that the gene product may be a secreted, soluble receptor and immune mediator.

Gene NO: 62 is expressed primarily in human fetal heart, meningima, and to a lesser extent in tonsils. This gene also is expressed in the breast cancer cell line MDA

10 36.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, lymphomas, leukemias, breast cancer and any immune system dysfunction, including those dysfunctions which involve natural killer cell activities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system or breast cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., heart, meningima, and tonsils and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the LAG3 gene (murine) indicates that the polynucleotides and polypeptides corresponding to Gene NO: 62 are useful for diagnostic and/or therapeutic modalities directed at abnormalities or disease states involving defective immune systems, preferably involving natural killer cell activity, as well as breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 195 as residues: Pro-10 to Trp-17, Cys-58 to Pro-67, Thr-76 to Glu-85, and Arg-93 to Asn-101.

**35 FEATURES OF PROTEIN ENCODED BY GENE NO: 63**

The translation product of Gene NO: 63 shares sequence homology with a *Caenorhabditis elegans* alpha-collagen gene (Clg), which is thought to be important in

organism development, as well as other collagen genes. Thus, based on sequence homology, polypeptides of this gene are expected to have activity similar to collagen, including involvement in organ development.

Gene NO: 63 is expressed primarily in human B-Cell Lymphoma, and to a lesser extent in human pituitary tissue. This gene has also demonstrated expression in keratinocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B-Cell Lymphoma, other lymphomas, leukemias, and other cancers, as well as disorders related to development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., tissue and/or cells of the immune system, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* alpha-collagen gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 63 are useful for development of diagnostic and/or therapeutic modalities directed at the detection and/or treatment of cancer, specifically B-Cell Lymphomas, leukemias, or diseases related to development.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 196 as residues: Thr-22 to Arg-27 and Ser-29 to Thr-39.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of Gene NO: 64 shares sequence homology with human extracellular molecule olfactomedin, which is thought to be important in the maintenance, growth, or differentiation of chemosensory cilia on the apical dendrites of olfactory neurons. Based on this sequence homology, it is likely that polypeptides of this gene have activity similar to the olfactomedin, particularly the differentiation or proliferation of neurons.

Gene NO: 64 is expressed primarily in fetal lung tissue.

- Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the lung as well as neural development, particularly of the
- 5 lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., lungs and cancerous and wounded tissues)
- 10 or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the olfactomedin family indicates that

15 polypeptides and polynucleotides corresponding to Gene NO: 64 are useful for the development of diagnostic and/or therapeutic modalities directed at detection and/or treatment of pulmonary disease states, e.g., cystic fibrosis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 197 as residues: Gly-17 to Gln-23, Gln-45 to Arg-50, Arg-56 to Lys-61, Glu-70 to

20 Leu-76, Asp-88 to Glu-93, Pro-117 to Met-131, Asp-161 to Glu-167, Arg-224 to Asn-237, Asp-302 to Trp-312, Pro-315 to Asn-320, and Thr-337 to Ser-341.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 65

- The translation product of Gene NO: 65 shares sequence homology with
- 25 *Saccharomyces cerevisiae* hypothetical protein YKL166 (Accession No. gi/687880) which is thought to be important in secretory and/or vesicular transport mechanisms. Based on this homology, it is likely that the gene product would have similar activity to YKL166, particularly secretory or transport mechanisms. Preferred polypeptide fragments of this gene include those fragments starting with the amino acid sequence
- 30 ISAARV (SEQ ID NO:271) . Other polypeptide fragments include the former fragment, which ends with the amino acid sequence PDVSEFMTRLF (SEQ ID NO:272). Further preferred fragments include those polypeptide fragments comprising the amino acid sequence FDPVRVDITSKGKMRAR (SEQ ID NO:273). Also preferred are polypeptide fragments having exogenous signal sequences fused to the polypeptide.
- 35 Gene No 65 is expressed primarily in placenta, testis, osteoclastoma and to a lesser extent in adrenal gland.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or diseases involving defects in protein secretion. Similarly, 5 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, cartilage and bone, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, testis, 10 adrenal gland, and osteoclastoma, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 The tissue distribution and homology to the yeast YKL1GG protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 65 are useful for the development of therapeutic and/or diagnostic modalities targeted at cancer or secretory anomalies, such as genetically caused secretory diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 20 198 as residues: Ser-18 to Ser-29 and Lys-53 to Arg-74.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 66

The translation product of Gene NO: 66 shares sequence homology with the human papilloma virus (HPV) E5 ORF region which is thought to be important as a 25 secreted growth factor. Although this is described as a viral gene product, it is believed to have several cellular secretory homologues. Therefore, based on the sequence similarity between the HPV E5 ORF and the translated product of this gene, this gene product is likely to have activity similar to HPV E5 ORF.

Gene NO: 66 is expressed primarily in activated T-Cells, monocytes, 30 cerebellum and to a lesser extent in infant brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or human papilloma virus infection. Similarly, 35 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of

this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, lymph tissue, monocytes, and T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Moreover, polynucleotides of this gene have been mapped to chromosome 1. Therefore, polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 1.

The tissue distribution and homology to human papilloma virus E5 region indicates that polypeptides and polynucleotides corresponding to Gene NO: 66 are useful for development of diagnostic and/or therapeutic modalities directed at the diagnosis and/or treatment of cancer and/or human papilloma virus infection (HPV).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 199 as residues: Asn-31 to Arg-36 and Leu-102 to Ser-112.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The translation product of Gene NO: 67 shares sequence homology with the 8hs20 protein precursor [*Mus musculus*] which is thought to be important in B-Cell mu chain assembly. (See, Accession No. PID/d1002996; Shiraswa, T., EMBO J. 12(5):1827-1834 (1993).) A polypeptide fragment starting at amino acid 53 is preferred, as well as 1-20 amino acid N-terminus and/or C-terminus deletions. Based on the sequence similarity between 8hs20 protein and the translation product of this gene, the two polypeptides are expected to share certain biological activities, particularly immunologic activities.

Gene NO: 67 is expressed primarily in human B-cells and to a lesser extent in Hodgkin's Lymphoma. It is also likely that the polypeptide will be expressed in B-cell specific cells, bone marrow, and spleen, as is observed with 8hs20.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, Hodgkin's Lymphoma, Common Variable Immunodeficiency, and/or other B-cell lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., bone

marrow, spleen, lymph tissue, and B-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5        The tissue distribution and homology to 8hs20 protein precursor [*Mus musculus*], indicates that polypeptides and polynucleotides corresponding to Gene NO: 67 are useful for therapeutic and/or diagnostic purposes, targeting Hodgkin's Lymphoma, B-cell lymphomas, Common Variable Immunodeficiency, or other
- 10      immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 200 as residues: Asp-51 to Trp-56, Arg-72 to Asp-85, and Gln-106 to Asp-112.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 68

- 15      Gene NO: 68 is expressed primarily in fetal liver/spleen, rhabdomyosarcoma, and to a lesser extent in 9 week-old early stage human embryo and bone marrow. Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
- 20      not limited to, rhabdomyosarcoma and other cancers, hematopoietic disorders, and immune dysfunction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or
- 25      lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, striated muscle, liver, spleen, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an
- 30      individual not having the disorder.

The tissue distribution indicates that the protein product of Gene NO: 68 is useful for diagnostic and/or therapeutic purposes directed to cancer, preferably rhabdomyosarcoma. Enhanced expression of this gene in fetal liver, spleen, and bone marrow indicates that this gene plays an active role in hematopoiesis. Polypeptides or polynucleotides of the present invention may therefore help modulate survival, proliferation, and/or differentiation of various hematopoietic lineages, including the hematopoietic stem cell. Thus, polynucleotides or polypeptides can be used treat

various hematopoietic disorders and influence the development and differentiation of blood cell lineages, including hematopoietic stem cell expansion. The polypeptide does contain a thioredoxin family active site at amino acids 64-82. Polypeptides comprising this thioredoxin active site are contemplated.

5

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 69**

Gene NO: 69 is expressed primarily in liver and kidney and to a lesser extent in macrophages, uterus, placenta, and testes.

- Therefore, polynucleotides or polypeptides of the invention are useful as  
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal disorders, neoplasms (e.g., soft tissue cancer, hepatocellular tumors), immune disorders, endocrine imbalances, and reproductive disorders.  
Similarly, polypeptides and antibodies directed to these polypeptides are useful in  
15 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic, urogenital, immune, and reproductive systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, kidney, uterus, placenta, testes, and macrophages and cancerous and  
20 wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution indicates that polypeptides and polynucleotides  
25 corresponding to Gene NO: 69 are useful for diagnosis and treatment of disorders in the hepatic, urogenital, immune, and reproductive systems.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 202 as residues: Arg-41 to Ser-50, Glu-138 to Asn-148, Ser-155 to Arg-172, Pro-219 to Glu-228.

30

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 70**

Gene NO: 70 is expressed primarily in the immune system, including macrophages, T-cells, and dendritic cells and to a lesser extent in fetal tissue.

- Therefore, polynucleotides or polypeptides of the invention are useful as  
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, inflammatory diseases, lymph node disorders, fetal

development, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems expression of this gene at 5 significantly higher or lower levels may routinely be detected in certain tissues and certain cell types (e.g., macrophages, T-cells, dendritic cells, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level 10 in healthy tissue or bodily fluid from an individual not having the disorder. There is some evidence that the polynucleotide is mapped to chromosome 19. Thus, the polynucleotide can be a marker for genetic analysis for chromosome 19.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 70 are useful for treatment, prophylaxis, and diagnosis of 15 immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. The polypeptides or polynucleotides of the present invention are also useful in the treatment, prophylaxis, and detection of thymus disorders, such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism. The expression observed 20 predominantly in hematopoietic cells also indicates that the polynucleotides or polypeptides are important in treating and/or detecting hematopoietic disorders, such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and 25 functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The polypeptides or polynucleotides are also useful to increase the proliferation of peripheral blood leukocytes, which can be used in the combat of a range of hematopoietic disorders, including immunodeficiency diseases, leukemia, and 30 septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 203 as residues: Thr-21 to Ser-27, Pro-33 to Ser-38, and Arg-73 to Lys-84.

64

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of Start Codon	AA of Signal Pep	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
1	HGCMID20	97901 02/26/97 209047 05/15/97	pSport1	11 1739	25 1658	54 54	134 134	1 1	28 28	29 29	466
2	HLDBG33	97898 02/26/97 209044 05/15/97	pCMV Sport 3.0	12 844	1 844	39 39	135 135	1 1	28 28	29 29	221
2	HLDBG33	97898 02/26/97 209044 05/15/97	pCMV Sport 3.0	81 795	1 795	10 434	10 10	204 204	1 1	29 29	34
3	HTGEW86	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	13 776	134 676	173 173	136 136	1 1	35 35	36 36	155
4	HKCSR70	97900 02/26/97 209046 05/15/97	pBluescript	14 1376	727 1343	202 202	137 137	1 1	20 20	21 21	232
4	HKCSR70	97900 02/26/97 209046 05/15/97	pBluescript	82 1324	741 741	1309 861	205 205	1 1	31 31	32 32	42

65

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO.: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
4	HETB187	209010 04/28/97 209085 05/29/97	Uni-ZAP XR	83	1494	1	1484	51	51	206	1	34
5	HTEAU17	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	15	502	1	502	143	143	138	1	33
6	HBMCY91	97897 02/26/97 209043 05/15/97	pBluescript	16	425	1	425	56	56	139	1	17
7	HSSGE07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	17	1316	1	1298	45	45	140	1	26
7	HSSGE07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	84	1285	1	1271	15	15	207	1	28
8	HBMBX59	97897 02/26/97 209043 05/15/97	pBluescript	18	436	87	384	157	157	141	1	21

66

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT Total Seq. No. X	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F		
9	HNGIT22	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	19	503	1	503	23	23	142	1	19	20	40
10	HERAD57	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	20	358	1	358	147	147	143	1	31	32	69
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	21	1926	573	1926	157	157	144	1	30	31	482
11	HCENJ40	97898* 02/26/97 209044 05/15/97	Uni-ZAP XR	85	394	1	394	166	166	208	1	20	21	23
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	86	1925	573	1925	157	157	209	1	30	31	482
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	87	1818	30	1298	1137	1137	210	1			12

67

67

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT of Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
12	HCSR90	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	22	1224	64	557	80	80	145	1	30
13	HBJFC03	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	23	694	1	694	181	181	146	1	39
13	HBJFC03	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	88	539	1	539	215	215	211	1	18
14	HSNBL85	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	24	796	405	796	1	1	147	1	30
14	HSNBL85	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	89	855	300	855	513	513	212	1	37
15	HTEBY26	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	25	662	205	653	77	77	148	1	30

68

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Total NT Seq.	5' NT of Clone Seq.	First AA of Signal Pep.	5' NT of Start Codon	AA SEQ ID NO: Y	First AA of Sig Pep.	Last AA of Sig Pep.	Predicted First AA of Secreted Portion	Last AA of OR F
15	HTEBY26	97899 02/26/97 209045	Uni-ZAP XR	90	628	198	625			275	213	1	31	32
16	HMABH07	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	26	1105	40	1105	88	88	149	1	18	19	164
16	HMABH07	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	91	1053	61	1009	79	79	214	1	22	23	229
17	HSKNY94	97899 02/26/97 209045 05/15/97	pBluescript	27	1017	1	1017	97	97	150	1	30	31	138
17	HSKNY94	97899 02/26/97 209045 05/15/97	pBluescript	93	2492	1	943	100	100	216	1	27	28	126
18	HMCDA67	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	28	391	1	391	169	169	151	1	29	30	57

69

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT Total Seq.	5' NT of Clone Seq.	3' NT of Start Codon	5' NT of AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
19	HOSFF45	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	29	1139	6	1139	109	109	152	1
19	HOSFF45	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	94	3058	1795	2847	1868	1868	217	1
20	HMJAA51	97899 02/26/97 209045 05/15/97	pSportI	30	465	1	370	47	47	153	1
20	HMJAA51	97899 02/26/97 209045 05/15/97	pSportI	95	1099	664	1000	669	669	218	1
21	HTEBF05	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	31	702	1	702	403	403	154	1
22	HTEAL31	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	32	1142	1	518	49	49	155	1

70

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT Total Seq.	3' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of First AA of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
22	HTEAL31	97899 02/26/97 209045	05/15/97 02/26/97 209045	Uni-ZAP XR 96	1580	23	422	32	219	1	47	48
23	HBMCT32	97899 02/26/97 209045	05/15/97 02/26/97 209045	pBluescript 33	928	1	928	48	156	1	27	28
23	HBMCT32	97899 02/26/97 209045	05/15/97 02/26/97 209045	pBluescript 97	678	72	593	89	220	1	27	28
24	HSKXE91	97899 02/26/97 209045	05/15/97 02/26/97 209045	pBluescript 34	773	1	773	39	39	157	1	22
24	HSKXE91	97899 02/26/97 209045	05/15/97 02/26/97 209045	pBluescript 98	1253	507	1253	507	507	221	1	16
25	HPWTB39	97899 02/26/97 209045	05/15/97 02/26/97 05/15/97	Uni-ZAP XR 35	453	1	453	40	40	158	1	25
												74

71

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of AA of Start Signal Codon	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
26	HTLEV12	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	36	459	1	459	25	25	159	1
27	HSPAF93	97900 02/26/97 209046 05/15/97	pSport1	37	509	1	509	1	1	160	1
27	HSPAF93	97900 02/26/97 209046 05/15/97	pSport1	99	447	1	447	7	7	222	1
28	HHFGL62	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	38	598	1	598	1	1	161	1
28	HHFGL62	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	100	611	37	611	17	17	223	1
29	HCEIU14	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	39	454	1	454	1	1	162	1

72

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z	NT SEQ ID NO: X	NT Total Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
29	HCEIU14	97900 02/26/97 209046	Uni-ZAP XR	101	609	176	609	237	237	224	1	14
30	HEBDA39	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	40	425	1	376	223	223	163	1	18
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	41	2471	141	2471	213	213	164	1	30
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	102	1770	47	1721	119	119	225	1	31
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	103	1832	96	1777	138	138	226	1	9
32	HAGBB70	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	42	2659	1172	2659	119	119	165	1	18
												103

73

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT Total NT Seq.	3' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of First AA of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
32	HAGBB70	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	104	2237	878	2237	1134	227	1		19
33	HETDG84	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	43	1635	100	1580	299	299	166	1	20
34	HTEGA81	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	44	780	19	717	10	10	167	1	23
34	HKGAJ40	209236 09/04/97	pSportI	105	1822	1	1023	272	272	228	1	23
34	HKMLK44	209084 05/29/97	pBluescript	106	1712	1	1669	168	168	229	1	21
35	HTXAK60	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	45	2378	1337	2378	1437	1437	168	1	30
35	HTXAK60	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	107	1969	1068	1892	989	989	230	1	23
												36

74

Gene No.	CDNA Clone ID	ATCC Deposit No.: Z	NT SEQ ID NO: X	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Total NT Seq.	5' NT of Clone Seq.	AA of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of QR F		
36	HMHB40	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	46	1772	69	1772			129	169	1	30	31	231
36	HMHB40	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	108	1734	65	1734	100	100	231	1	29	30	80	
37	HFVGS85	97901 02/26/97 209047 05/15/97	pBluescript	47	1107	70	1107	83	83	170	1	30	31	71	
38	HERAH81	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	48	805	167	764	167	167	171	1	23	24	64	
39	HMSEU04	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	49	1408	131	1258	364	364	172	1	22	23	74	
40	HNED157	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	50	1813	1	1184	2	2	173	1	1	2	333	

75

75

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT Total Seq.	5' NT of Clone Seq.	3' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F		
41	HNTME13	97901 02/26/97 209047 05/15/97	pSport1	51	2070	74	2070	142	142	174	1	20	21	195
41	HNTME13	97901 02/26/97 209047 05/15/97	pSport1	109	2003	15	1957	68	68	232	1	22	23	300
42	HSXBT25	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	52	1426	1	1426	158	158	175	1	23	26	264
42	HSXBT25	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	110	1320	80	1311	41	41	233	1	29	30	312
43	HSXCK41	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	53	1720	1	1720	161	161	176	1	22	23	137
43	HSXCK41	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	111	1962	299	1962	566	566	234	1	33	34	47

76

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT Total Seq. No: X	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F	
44	HE8CJ26	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 1117	1	1107	218	218	177	1	25	26	178
44	HE8CJ26	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 112	1785	30	1087		225	235	1	23	24
45	HTTDS54	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 55	1903	1	1903	119	119	178	1	31	32
45	HTTDS54	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 113	1842	1	1832	80	80	236	1	36	37
46	HLHDY31	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 56	1869	133	1838	124	124	179	1	24	25
46	HLHDY31	97902 02/26/97 209048	05/15/97 209048	Uni-ZAP XR 114	1960	90	1960	165	165	237	1	24	25

77

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Signal Pep.	5' NT of AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
47	HMCBP63	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	57	1259	320	1010	352	180	1	26	27
48	HEMGE83	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	58	1186	33	557	12	12	181	1	18
49	HHSDC22	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	59	428	1	304	172	172	182	1	34
50	HHSDZ57	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	60	501	1	501	40	40	183	1	62
50	HHSDZ57	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	115	536	73	536	73	73	238	1	22
52	HMMAB12	97903 02/26/97 209049 05/15/97	pSport1	62	595	1	595	308	308	185	1	29
												30
												42

78

Gene No.	cDNA Clone ID	ATCC Deposit No. Z and Date	NT SEQ ID NO: X	NT Total Seq. NT Seq.	3' NT of Clone Seq.	5' NT of Clone Seq.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted Secreted Portion	Last AA of F			
52	HMMAB12	97903 02/26/97 209049 05/15/97	pSport1	118	453	1	453	198	198	241	1	26	27	27
53	HSKDW02	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	63	1478	40	1436	176	176	186	1	39	40	58
53	HSKDW02	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	119	2016	211	1957	317	317	242	1	25	26	57
54	HETGL41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	64	2033	1	2033	30	30	187	1	30	31	187
54	HETGL41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	120	2136	110	2134	296	296	243	1	23	24	122
55	HODAZ50	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	65	440	1	440	1	1	188	1	26	27	145

79

Gene No.	cDNA Clone ID	ATCC Deposit No.:Z and Date	NT SEQ ID NO: X	NT Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F		
55	HODAZ50	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	121	219	1	219			1	244	1	10	11	72
56	HSDGE59	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	66	3301	349	1478	341	341	189	1	30	31	83	
57	HE6ES13	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	67	1535	1	1535	331	331	190	1	26	27	57	
57	HE6ES13	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	122	1686	239	1678								
58	HSSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	68	1244	402	1244	57	57	191	1	30	31	310	
58	HSSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	123	1211	1	1211	80	80	246	1	30	31	338	

80

80

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT Total NT Seq.	3' NT of Clone Seq.	5' NT of Clone Seq.	S' NT of First AA of Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of Secreted Portion	OR F
58	HSSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	124	1804	402	1526	501	247	1			17
59	HRDEV41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	69	1292	1	1278	70	70	192	1	28	29
59	HRDEV41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	125	1282	31	1088	70	70	248	1	21	22
60	HILCJ01	97903 02/26/97 209049 05/15/97	pBluescript SK-	70	1031	498	1031	536	536	193	1	30	31
61	HSATP28	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	71	855	178	855	187	187	194	1	28	29
62	HFGL41	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	72	1274	58	1274	118	118	195	1	42	43
													101

81

81

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID NO: X	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of AA of Start Signal Codon	5' NT of AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
62	HHFGL41	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	126	1296	88	1237	133	249	1	39	40
63	HBJEM49	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	73	688	1	688	173	173	196	1	18
63	HBJEM49	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	127	737	1	737	174	174	250	1	20
64	HSLDJ95	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	74	1890	1	1890	112	112	197	1	21
64	HSLDJ95	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	128	1925	1	1829	87	87	251	1	23
65	HSREG44	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	75	1133	408	1133	531	531	198	1	18
												73

82

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	NT SEQ ID: X	NT Total Seq. NO: X	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA of Signal Pep.	AA SEQ ID: NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
66	HTXCT40	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	76	585	1	585	1	199	1	69	70
66	HTXCT40	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	129	2713	2023	2713	2133	252	1	39	40
67	HRGDF73	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	77	577	1	577	51	200	1	23	24
68	HRDBF52	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	78	2278	1458	1935	25	201	1	23	24
68	HRDBF52	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	130	1011	479	1011	701	253	1	20	21
68	HKMND45	97976 04/04/97	pBluescript	131	2278	1	1929	25	254	1	27	28
69	HPEBD70	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	79	1143	601	1097	95	202	1	6	7

83

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	NT SEQ ID NO: X	NT SEQ ID NO: X	5' NT of Total NT Seq.	3' NT of Clone Seq.	5' NT of Start Signal Pep.	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F		
69	HPEBD70	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	132	1088	535	1043	588	588	255	1	27	28	52
70	HMCAB89	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	80	557	1	557	132	132	203	1	25	26	92

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification , such as multiple histidine residues, or an additional sequence for stability during recombinant production.

- 5        The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988).  
Polypeptides of the invention also can be purified from natural or recombinant sources  
10      10 using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

### Signal Sequences

- Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch,  
15      Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1  
20      indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

- In the present case, the deduced amino acid sequence of the secreted polypeptide  
25      was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results  
30      shown in Table 1.

- As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +  
35      or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

- 15 "Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994); SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).
- 20 Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park,
- 25 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981).)

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

- 5 A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 10 (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB search of a DNA sequence to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch 15 Penalty=1, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group 20 Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone, means that the polynucleotide is identical to a sequence contained in 25 SEQ ID NO:X or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:X or the deposited clone, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:X or the cDNA can be deleted, 30 inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity

will encode a polypeptide identical to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

5       Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988  
10      (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

15      Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible  
20      amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

25      Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

30      Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as  
35      have little effect on activity. For example, guidance concerning how to make

phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

#### Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, and 701 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, and 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about"

includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred.

Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

### Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')<sub>2</sub> fragments) which are capable of specifically binding to protein. Fab and F(ab')<sub>2</sub> fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

### Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

5 Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the  
10 polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of  
15 immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86  
20 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion  
25 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,  
30 would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol.  
35 Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In

preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein.

5 Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the claimed invention.

10

#### Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

15

20 The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

25

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

30

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS,

293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

### Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes

5 known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention

10 can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic

15 cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping

20 strategies that can be used include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence *in situ* hybridization (FISH) of a metaphase chromosomal spread. This

25 technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

30 For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross

35 hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage

analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, *Mendelian Inheritance in Man* (available on line through Johns Hopkins University Welch Medical Library).) Assuming 1 megabase mapping resolution and one gene per

- 5 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or

- 10 translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the 15 mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, 20 chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

- In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred 25 polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC 30 Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

- 35 Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the

present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute

- 5 biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" 10 which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

- The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an 15 individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely 20 small tissue samples.

- Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from 25 polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the 30 present invention can be used as polymorphic markers for forensic purposes.

- There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present 35 invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers 5 for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

### Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The 10 following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-15 3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and 20 technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of protein include those detectable by X-radiography, NMR or ESR. For X- 25 radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

30 A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the 35 subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20

5 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

10 Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

15 Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

20 Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

25 At the very least, the polypeptides of the present invention could be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

### Biological Activities

35 The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules

may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

### Immune Activity

5        A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells  
10      from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

15      A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic  
20      cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency  
25      (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood  
30      coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks  
35      (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from

inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

5 Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, 10 Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune 15 inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

20 A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The 25 administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

30 Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute 35 rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

### Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

### Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, 5 Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., 10 Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), 15 pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

20 Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Nocardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, 25 Blastomycosis, Bordetella, Brorrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocytoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, 30 Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, 35 respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria,

Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect 5 any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidioidomycosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, 10 Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide 15 of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide 20 of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

### Regeneration

A polynucleotide or polypeptide of the present invention can be used to 25 differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal 30 disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal 35 or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

20

#### Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat

disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

### **Binding Activity**

5 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or  
10 small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural  
15 receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

20 Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

25 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

30 Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

35 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The

antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

15

#### Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

20

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

25

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

30

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3'

10 Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

15 Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

20 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

25 Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

30 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

35 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or 5 of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said 10 cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the 15 ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone. 20

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

25 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

30 A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method 35 comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining

whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the

amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an

amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5 Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

10 Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

15 Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

20 Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

25 In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

30 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least

90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 5 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

10 Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the 15 ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the 20 recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said 25 polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human 30 cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased 35 level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated

polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of 5 illustration and are not intended as limiting.

### Examples

#### Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

10        Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the 15 related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
	Lambda Zap	pBluescript (pBS)
20	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSport 2.0	pCMVSport 2.0
25	pCMVSport 3.0	pCMVSport 3.0
	pCR®2.1	pCR®2.1

      Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 30: 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS-. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which

are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from 5 Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 10 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: 15 (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing 20 the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

25 Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized 30 using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ - $\gamma$ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as 35 XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above.

The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 5 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 10 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM 15 MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

20 Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the 25 missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population 30 of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged 35 RNA which may interfere with the later RNA ligase step. The phosphatase should then

be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

- 5        This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that  
10      the 5' end sequence belongs to the desired gene.

**Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

- A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR  
15      using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

**Example 3: Tissue Distribution of Polypeptide**

- Tissue distribution of mRNA expression of polynucleotides of the present  
20      invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P<sup>32</sup> using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.),  
25      according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

- Multiple Tissue Northern (MTN) blots containing various human tissues (H) or  
human immune system tissues (IM) (Clontech) are examined with the labeled probe  
using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's  
30      protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

**Example 4: Chromosomal Mapping of the Polynucleotides**

- An oligonucleotide primer set is designed according to the sequence at the 5'  
35      end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This

primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual 5 chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

10 **Example 5: Bacterial Expression of a Polypeptide**

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as 15 BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp<sup>r</sup>), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites. 20

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses 25 the lacI repressor and also confers kanamycin resistance (Kan<sup>r</sup>). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). 30 The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D. <sup>600</sup>) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number XXXXXX.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or

Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

5

**Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

10 Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50  
15 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

15 The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by  
20 centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

25 The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

30 Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive

Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

5       Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium  
10      acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant  $A_{280}$  monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

15      The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5  $\mu$ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

20

**Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System**

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong  
25      polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the  
30      same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such  
35      as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription,

translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the  
5 AUG initiation codon and the naturally associated leader sequence identified in Table 1,  
is amplified using the PCR protocol described in Example 1. If the naturally occurring  
signal sequence is used to produce the secreted protein, the pA2 vector does not need a  
second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a  
baculovirus leader sequence, using the standard methods described in Summers et al.,  
10 "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures,"  
Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially  
available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested  
with appropriate restriction enzymes and again purified on a 1% agarose gel.

15 The plasmid is digested with the corresponding restriction enzymes and  
optionally, can be dephosphorylated using calf intestinal phosphatase, using routine  
procedures known in the art. The DNA is then isolated from a 1% agarose gel using a  
commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4  
20 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue  
(Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation  
mixture and spread on culture plates. Bacteria containing the plasmid are identified by  
digesting DNA from individual colonies and analyzing the digestion product by gel  
electrophoresis. The sequence of the cloned fragment is confirmed by DNA  
25 sequencing.

Five µg of a plasmid containing the polynucleotide is co-transfected with 1.0 µg  
of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus  
DNA", Pharmingen, San Diego, CA), using the lipofection method described by  
Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One µg of  
30 BaculoGold™ virus DNA and 5 µg of the plasmid are mixed in a sterile well of a  
microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies  
Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are  
added, mixed and incubated for 15 minutes at room temperature. Then the transfection  
mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm  
35 tissue culture plate with 1 ml Grace's medium without serum. The plate is then  
incubated for 5 hours at 27° C. The transfection solution is then removed from the plate

and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life

5 Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a  
10 micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

15 To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine  
20 (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of <sup>35</sup>S-methionine and 5 µCi <sup>35</sup>S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

25 Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

#### **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates  
30 the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from  
35 Retroviruses, e.g., RSV, HTLV, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used 5 include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the 10 polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the 15 encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992)). Using these markers, the 20 mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the 25 expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the 30 cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate 35 restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the

secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

#### Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., *Nature* 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the

activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which 5 outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an 10 expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with 15 BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally 20 occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCCAAATCTTCTGACAAAACTCACACATGCCACC GTGCC  
25 CAGCACCTGAATTGAGGGTGACCGTCAGTCTCCTCTTCCCCC AAAACC  
CAAGGACACCCCTCATGATCTCCGGACTCCTGAGGTACATGCGTGGTGGT  
GGACGTAAGCCACGAAGACCCCTGAGGTCAAGTTCACTGGTACGTGGACG  
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGGAGGAGCAGTACAAC  
AGCACGTACCGTGTGGTCAGCGTCCCTCACCGTCTGCACCAGGACTGGCTG  
30 AATGGCAAGGAGTACAAGTGAAGGTCTCCAACAAAGCCCTCCAAACCCCC  
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCCACAGGT  
GTACACCCCTGCCCTCATCCGGATGAGCTGACCAAGAACCAAGGT CAGCCT  
GACCTGCCTGGTCAAAGGCTCTATCCAAGCGACATGCCGTGGAGTGGGA  
GAGCAATGGGCAGCCGGAGAACAAACTACAAGACCACGCCCTCCCGT GCTGG  
35 ACTCCGACGGCTCCTCTCCTACAGCAAGCTCACCGTGGACAAGAGCA  
GGTGGCAGCAGGGAACGTCTCATGCTCCGTGATGCATGAGGCTCTGC

ACAAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC  
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

**Example 10: Production of an Antibody from a Polypeptide**

- 5       The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants.
- 10      Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 15 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in 20 any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma 25 cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells 30 obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is 35 possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a

mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and 5 can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain 10 (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using 15 genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 20 8702671; Boulian et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

**Example 11: Production Of Secreted Protein For High-Throughput Screening Assays**

25

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution 30 (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The 35 PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

5       The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a  
10      multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

15       Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel,  
20      adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (see below) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock  
25      solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours  
30      depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays  
35      described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other

proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

### 5 *HGS-CHO-5 medium formulation:*

#### Inorganic Salts

CaCl <sub>2</sub> (anhyd)	116.6 mg/L
CuSO <sub>4</sub> ·5H <sub>2</sub> O	0.00130
Fe(NO <sub>3</sub> ) <sub>3</sub> ·9H <sub>2</sub> O	0.050
FeSO <sub>4</sub> ·7H <sub>2</sub> O	0.417
KCl	311.80
MgCl <sub>2</sub>	28.64
MgSO <sub>4</sub>	48.84
NaCl	6995.50
NaHCO <sub>3</sub>	2400.0
NaH <sub>2</sub> PO <sub>4</sub> ·H <sub>2</sub> O	62.50
Na <sub>2</sub> HPO <sub>4</sub>	71.02
ZnSO <sub>4</sub> ·7H <sub>2</sub> O	.4320

#### Lipids

Arachidonic Acid	.002 mg/L
Cholesterol	1.022
DL-alpha-Tocopherol-Acetate	.070
Linoleic Acid	0.0520
Linolenic Acid	0.010
Myristic Acid	0.010
Oleic Acid	0.010
Palmitic Acid	0.010
Palmitic Acid	0.010
Pluronic F-68	100
Stearic Acid	0.010
Tween 80	2.20

### 10 Carbon Source

D-Glucose	4551 mg/L
-----------	-----------

#### Amino Acids

L- Alanine	130.85 mg/ml
L-Arginine-HCL	147.50
L-Asparagine-H <sub>2</sub> O	7.50

L-Aspartic Acid	6.65
L-Cystine-2HCL-H <sub>2</sub> O	29.56
L-Cystine-2HCL	31.29
L-Glutamic Acid	7.35
L-Glutamine	365.0
Glycine	18.75
L-Histidine-HCL-H <sub>2</sub> O	52.48
L-Isoleucine	106.97
L-Leucine	111.45
L-Lysine HCL	163.75
L-Methionine	32.34
L-Phenylalanine	68.48
L-Proline	40.0
L-Serine	26.25
L-Threonine	101.05
L-Tryptophan	19.22
L-Tyrosine-2Na-2H <sub>2</sub> O	91.79
L-Valine	99.65

### Vitamins

Biotin	0.0035 mg/L
D-Ca Pantothenate	3.24
Choline Chloride	11.78
Folic Acid	4.65
i-Inositol	15.60
Niacinamide	3.02
Pyridoxal HCL	3.00
Pyridoxine HCL	0.031
Riboflavin	0.319
Thiamine HCL	3.17
Thymidine	0.365
Vitamin B <sub>12</sub>	0.680

### Other Components

HEPES Buffer	25 mM
Na Hypoxanthine	2.39 mg/L
Lipoic Acid	0.105
Sodium Putrescine-2HCL	0.081
Sodium Pyruvate	55.0
Sodium Selenite	0.0067
Ethanolamine	20uM
Ferric Citrate	0.122
Methyl-B-Cyclodextrin complexed with Linoleic Acid	41.70

Methyl-B-Cyclodextrin complexed with Oleic Acid	33.33
Methyl-B-Cyclodextrin complexed with Retinal Acetate	10

*Adjust osmolarity to 327 mOsm*

**Example 12: Construction of GAS Reporter Construct**

5 One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

10 GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with 15 IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") 20 family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 25 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN- $\alpha$ , IFN- $\gamma$ , and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a 30 WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of 5 the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	<u>Ligand</u>	<u>tyk2</u>	<u>JAKs</u>			<u>STATS</u>	<u>GAS(elements) or ISRE</u>	
			<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>			
<b>IFN family</b>								
5	IFN-a/B	+	+	-	-	1,2,3	ISRE	
	IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)	
	Il-10	+	?	?	-	1,3		
<b>gp130 family</b>								
10	IL-6 (Pleiotropic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)	
	Il-11(Pleiotropic)	?	+	?	?	1,3		
	OnM(Pleiotropic)	?	+	+	?	1,3		
	LIF(Pleiotropic)	?	+	+	?	1,3		
	CNTF(Pleiotropic)	-/+	+	+	?	1,3		
15	G-CSF(Pleiotropic)	?	+	?	?	1,3		
	IL-12(Pleiotropic)	+	-	+	+	1,3		
<b>g-C family</b>								
20	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS	
	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)	
	IL-7 (lymphocytes)	-	+	-	+	5	GAS	
	IL-9 (lymphocytes)	-	+	-	+	5	GAS	
	IL-13 (lymphocyte)	-	+	?	?	6	GAS	
	IL-15	?	+	?	+	5	GAS	
25	<b>gp140 family</b>							
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)	
	IL-5 (myeloid)	-	-	+	-	5	GAS	
	GM-CSF (myeloid)	-	-	+	-	5	GAS	
30	<b>Growth hormone family</b>							
	GH	?	-	+	-	5		
	PRL	?	+/-	+	-	1,3,5		
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)	
35	<b>Receptor Tyrosine Kinases</b>							
	EGF	?	+	+	-	1,3	GAS (IRF1)	
	PDGF	?	+	+	-	1,3		
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)	
40								

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to

- 5 bind STATs upon induction with a range of cytokines (Rothman et al., *Immunity* 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:  
5':GCGCCTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCG

10 AAATGATTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

- PCR amplification is performed using the SV40 promoter template present in  
15 the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTCCCCGAAATCTAGATTCCCCGAAATGATTCCCCGAAATG  
20 ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCC  
CTAACTCCGCCATCCGCCCTAACTCCGCCAGTCCCCGCCATTCTCCGC  
CCCATGGGTACTAATTTTTATTATGCAGAGGCCGAGGCCCTCGGC  
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTGAGGCTAGGCTT  
TGCAAAAGCTT:3' (SEQ ID NO:5)

- 25 With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase,  
30 alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

- The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

**Example 13: High-Throughput Screening Assay for T-cell Activity.**

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

**Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells.

- 5 Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jak-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfet U937 cells with the GAS/SEAP/Neo construct produced 10 in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

15 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 1 mM  $\text{MgCl}_2$ , and 675 uM  $\text{CaCl}_2$ . Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then 20 resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting  $1 \times 10^8$  cells (this is enough for ten 96-well 25 plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200 ul cells per well in the 96-well plate (or  $1 \times 10^5$  cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma 30 can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

**Example 15: High-Throughput Screening Assay Identifying Neuronal Activity**

- 35

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfected PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCGG -3' (SEQ ID NO:6)  
5' GCGAAGCTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes Xhol/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

5 The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

10 Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to  $1 \times 10^5$  cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ml of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

15

**Example 16: High-Throughput Screening Assay for T-cell Activity**

NF- $\kappa$ B (Nuclear Factor  $\kappa$ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by 20 expression of certain viral gene products. As a transcription factor, NF- $\kappa$ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- $\kappa$ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF-  $\kappa$ B is retained in the cytoplasm with I- $\kappa$ B (Inhibitor  $\kappa$ B). However, upon stimulation, I-  $\kappa$ B is phosphorylated and degraded, causing NF-  $\kappa$ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF-  $\kappa$ B include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

30 Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- $\kappa$ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- $\kappa$ B would be useful in treating

diseases. For example, inhibitors of NF- $\kappa$ B could be used to treat those diseases related to the acute or chronic activation of NF- $\kappa$ B, such as rheumatoid arthritis.

To construct a vector containing the NF- $\kappa$ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- $\kappa$ B binding site (GGGGACTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:  
5' GCGGCCTCGAGGGACTTCCCGGGACTTCCGGGACTTCCGGAC  
TTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:  
10 5' GCGGCAAGCTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

15 Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5' CTCGAGGGACTTCCCGGGACTTCCGGGACTTCCGGGACTTCC  
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCGCCCTAACTCCGCCA  
20 TCCCGCCCTAACTCCGCCAGTCCGCCATTCTCCGCCCATGGCTGACT  
AATTTTTTATTATGCAGAGGCCAGGCCCTCGGCCTTGAGCTATT  
CAGAAGTAGTGAGGAGGCTTTTGAGGCCCTAGGCTTTGCAAAAAGCTT:  
3' (SEQ ID NO:10)

25 Next, replace the SV40 minimal promoter element present in the pSEAP2- promoter plasmid (Clontech) with this NF- $\kappa$ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- $\kappa$ B/SV40/SEAP cassette is removed from the above NF- $\kappa$ B/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- $\kappa$ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF- $\kappa$ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

5

**Example 17: Assay for SEAP Activity**

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the 10 following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

15

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15  $\mu$ l of 2.5x dilution buffer into Optiplates containing 35  $\mu$ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

20

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50  $\mu$ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50  $\mu$ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

25 Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

**Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

**Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

5 For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10<sup>6</sup> cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10<sup>6</sup> cells/ml, and dispensed into a microplate, 100 10 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

15 To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca<sup>++</sup> concentration.

20

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

30

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

5 Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodynne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr  
10 with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of  
15 alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodynne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of  
20 Loprodynne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 2 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub>  
25 and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum  
30 manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

35 Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a

biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

5       The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sub>2+</sub> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the  
10 components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

15      Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as  
20 above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of  
25 tyrosine kinase activity.

**Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase,  
30 Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other  
35

phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyn filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

25

**Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide**

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

35 PCR products is then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTHERM Polymerase. (Epicentre Technologies).

The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

- PCR products is cloned into T-tailed vectors as described in Holton, T.A. and 5 Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals is identified by mutations not present in unaffected individuals.

- Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated 10 according to Example 2 are nick-translated with digoxigenin-deoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

- Chromosomes are counterstained with 4,6-diamino-2-phenylidole and 15 propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. 20 et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovation Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated 25 disease.

**Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample**

- A polypeptide of the present invention can be detected in a biological sample, 30 and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

- For example, antibody-sandwich ELISAs are used to detect soluble 35 polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method

described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

#### Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes 5 of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules.

10 Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric 15 acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; 20 U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is 25 formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are 30 known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood 35 of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as 5 ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, 10 manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of 15 about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed 20 into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials 25 are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical 30 compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

**Example 24: Method of Treating Decreased Levels of the Polypeptide**

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form.

- 5 Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

**Example 25: Method of Treating Increased Levels of the Polypeptide**

- 15 Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 20 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

**Example 26: Method of Treatment Using Gene Therapy**

- 25 One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is 30 turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin, is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

5 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

10 The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to 15 transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

20 The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

25 Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is 30 required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is being produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

- 5        The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

## (1) GENERAL INFORMATION:

(i) APPLICANTS: Human Genome Sciences, Inc. et al.

(ii) TITLE OF INVENTION: 70 Human Secreted Proteins

5 (iii) NUMBER OF SEQUENCES: 273

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Human Genome Sciences, Inc.

(B) STREET: 9410 Key West Avenue

(C) CITY: Rockville

10 (D) STATE: Maryland

(E) COUNTRY: USA

(F) ZIP: 20850

## (v) COMPUTER READABLE FORM:

15 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

(B) COMPUTER: HP Vectra 486/33

(C) OPERATING SYSTEM: MSDOS version 6.2

(D) SOFTWARE: ASCII Text

## (vi) CURRENT APPLICATION DATA:

20 (A) APPLICATION NUMBER:

(B) FILING DATE: March 6, 1998

(C) CLASSIFICATION:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

25 (B) FILING DATE:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: A. Anders Brookes

(B) REGISTRATION NUMBER: 36,373

(C) REFERENCE/DOCKET NUMBER: PS001PCT

## (vi) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (301) 309-8504
- (B) TELEFAX: (301) 309-8439

## 5 (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

10 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGGATCCGGA	GCCCCAAATCT	TCTGACAAAA	CTCACACATG	CCCACCGTGC	CCAGCACCTG	60
AATTCCGAGGG	TGCACCGTCA	GTCTTCCTCT	TCCCCCCAAA	ACCCAAGGAC	ACCCTCATGA	120
15	TCTCCCGGAC	TCCTGAGGTC	ACATGCGTGG	TGGTGGACGT	AAGCCACGAA	GACCCCTGAGG
	TCAAGTTCAA	CTGGTACGTG	GACGGCGTGG	AGGTGCATAA	TGCCAAGACA	AAGCCGCGGG
	AGGAGGAGTA	CAACAGCACG	TACCGTGTGG	TCAGCGTCCT	CACCGTCCTG	CACCAAGACT
	GGCTGAATGG	CAAGGAGTAC	AAAGTGCAGG	TCTCCAACAA	AGCCCTCCCA	ACCCCCATCG
	AGAAAACCAT	CTCCAAAGCC	AAAGGGCAGC	CCCGAGAAC	ACAGGTGTAC	ACCCCTGCC
20	CATCCCGGGA	TGAGCTGACC	AAGAACCGAG	TCAGCCTGAC	CTGCCTGGTC	AAAGGCTTCT
	ATCCAAGCGA	CATCGCCGTG	GAGTGGGAGA	GCAATGGCA	GCCGGAGAAC	AACTACAAGA
	CCACGCCCTCC	CGTGCTGGAC	TCCGACGGCT	CCTTCTTCCT	CTACAGCAAG	CTCACCGTGG
	ACAAGAGCAG	GTTGGCAGCAG	GGGAACGTCT	TCTCATGCTC	CGTGATGCAT	GAGGCTCTGC
	ACAACCACTA	CACCGAGAAG	AGCCTCTCCC	TGTCTCCGGG	TAAATGAGTG	CGACGGCCGC
25	GACTCTAGAG	GAT				733

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Trp Ser Xaa Trp Ser

5 1 5

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 86 base pairs

10 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

15 GCGCCTCGAG ATTTCCCCGA AATCTAGATT TCCCCGAAAT GATTTCCCCG AAATGATTTC 60  
CCCGAAATAT CTGCCATCTC AATTTAG 86

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

25 GCGGCAAGCT TTTTGCAAAG CCTAGGC 27

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 271 base pairs

30 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

	CTCGAGATT CCCGAAATC TAGATTCCC CGAAATGATT TCCCCGAAT GATTTCCCCG	60
5	AAATATCTGC CATCTCAATT AGTCAGCAAC CATACTCCG CCCCTAACTC CGCCCATCCC	120
	GCCCCCTAACT CCGCCCAAGTT CGGCCCATTC TCCGGCCAT GGCTGACTAA TTTTTTTTAT	180
	TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTG CAGAAGTAGT GAGGAGGCCT	240
	TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T	271

10 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGCTCGAGG GATGACAGCG ATAGAACCCC GG	32
-------------------------------------	----

20 (2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCGAAGCTTC GCGACTCCCC GGATCCGCCT C	31
------------------------------------	----

30 (2) INFORMATION FOR SEQ ID NO: 8:

163

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGGACTTTC CC

12

## 10 (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GCGGCCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCGGGACT TTCCATCCTG

60

CCATCTCAAT TAG

73

20

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 256 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

30 CTCGAGGGGA CTTCAGGGGG GACTTTCCGG GGACTTTCCG GGACTTTCCA TCTGCCATCT

60

CAATTAGTCA GCAACCATAG TCCCGCCCT AACTCCGCCA ATCCCGCCC TAACTCCGCC	120
CAGTTCCGCC CAPPCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA	180
GGCCGCCTCG GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTG GAGGCCTAGG	240
CTTTGCAAA AAGCTT	256

5

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1739 base pairs

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCGCTCCCGA GGCGCGGGGA CCTGCAGAGA GGACAGCCGG CCTGCCCGG GACATGCCGC	60
15 CCCAGGAGCT CCCCAGGCTC GCGTTCCCGT TGCTGCTGTT GCTGTTGCTG CTGCTGCCGC	120
CGCCGCCGTG CCCTGCCAC AGGCCACGC GTTTCGACCC CACCTGGGAG TCCCTGGACG	180
CCCGCCAGCT GCCCGCGTGG TTIGACCAGG CCAAGTTGG CATCTTCATC CACTGGGAG	240
TGTTTCCGT GCCCAGCTTC GGTAGCGAGT GGTTCTGGTG GTATTGGAA AAGGAAAAGA	300
TACCGAAGTA TGTGGAATTG ATGAAAGATA ATTACCCCTCC TARTTTCAA TATGAAGATT	360
20 TTGGACCACT ATTTACAGCA AAATTTTTA ATGCCAACCA RTGGCARAT ATTTTYCAGG	420
CCTCTGGTGC CAAATACATT GTCTTAACCTT CCAAACATCA TGAAGGCTTT ACCTTGTGG	480
GGTCAGAATA TTCGTGGAAC TGGAAATGCCA TAGATGAGGG GCCCAAGAGG GACATTGTCA	540
AGGAACCTGA GGTAGCCATT AGGAACAGAA CTGACCTGCG TTTTGGACTG TACTTATTCCC	600
TTTTGAATG GTTTCATCCG CTCTTCCCTG AGGATGAATC CAGTTCATTC CATAAGCGGC	660
25 AATTTCCAGT TTCTAAAGACA TTGCCAGAGC TCTATGAGTT AGTGAACAAAC TATCAGCCTG	720
AGGTTCTGTG GTCGGATGGT GACGGAGGAG CACCGGATCA ATACTGGAAC ANCACAGGCT	780
TCTTGGCTG GTTATATAAT GAAAGCCAG TTCCGGGCAC AGTAGTCACC AATGATCGTT	840
GGGGAGCTGG TAGCATCTGT AAGCATGGTG GCTTCTATAC CTGCAGTGAT CGTTATAACC	900
CAGGACATCT TTGCCCCACAT AAATGGGAAA ACTGCATGAC AATAGACAAA CTGCTCTGGG	960
30 GCTATAGGAG GGAAGCTGGA ATCTCTGACT ATCTTACAAT TGAAGAATTG GTGAAGCAAC	1020

	TITGTAGAGAC AGTTTCATGT GGAGGAAATC TTTTGATGAA TATTGGGCC ACACTAGATG	1080
	GCACCATTTC TGTTAGTTTT GAGGAGCGAC TGAGGCCAAT GGGGTCTGG CTAAGAGTCA	1140
	ATGGAGAAC TATTTATGAA ACCCATACCT GGCGATCCC GAATGACACT GTCACCCAG	1200
	ATGTGTGGTA CACATCCAAG CCTAAAGAAA AATTAGTCTA TGCCATTPTT CTTAAATGGC	1260
5	CCACATCAGG ACAGCTGTT CTTGGCCATC CCAAAGCTAT TCTGGGGCA ACAGAGGTGA	1320
	AACTACTGGG CCATGGACAG CCACCTTAACG GGATTTCTTT GGAGCAAAAT GGCATTATGG	1380
	TAGAACTGCC ACAGCTAACC ATTCACTCAGA TGCCGTGTAATGGGGCTGG GCTCTAGCCC	1440
	TRACTAATGT GATCTAAAGT GCAGCAGAGT GGCTGATGCT GCAAGTTATG TCTAAGGCTA	1500
	GGAACTATCA GGTGTCTATA ATTGTAGCAC ATGGAGAAAG CAAATGTAAA ACTGGATAAG	1560
10	AAAATTATTT TGGCAGTTCA GCCCTTTCCC TTTTCCCAC TAAATTTTT CTTAAATTAC	1620
	CCATGTAACC ATTCTAACTC TCCAGTGCAC TTTGCCATTA AAGTCTCTTC ACATTGAAAA	1680
	AAAAAAAAAAA AAAAACCCCG GGGGGGGGGC CCGGGNACCC CATTCGCC NTAAAGGGG	1739

## (2) INFORMATION FOR SEQ ID NO: 12:

## 15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 844 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

	GGCCCCCTGGG CCCGAGGGGC TGGAGCCGGG CCGGGCGAT GTGGAGCGCG GGCCGCGCG	60
	GGGCTGCCCTG GCCGGTGCTG TTGGGGCTGC TGCTGGCGCT GTTAGTGCG GCGGGTGGTG	120
	CCGCCAAGAC CGGTGCGGAG CTCGTGACCT GCGGGTCGGT GCTGAAGCTG CTCAATACGC	180
	ACCACCGCGT GCGGCTGCAC TCGCACGACA TCAAATACGG ATCCGGCAGC GGCCAGCAAT	240
25	CGGTGACCGG CGTAGAGGCG TCGGACGACG CCAATAGCTA CTGGCGGATC CGCGGCGGCT	300
	CGGAGGGCGG GTGCCGCCGC GGGTCCCCGG TGCGCTGCCG GCAGGGCGTG AGGCTCACGC	360
	ATGTGCTTAC GGGCAAGAAC CTGCACACGC ACCACTTCCC GTCGCCGCTG TCCAACAACC	420
	AGGAGGTGAG TGCCTTGGG GAAGACGGCG AGGGCGACGA CCTGGACCTA TGGACAGTGC	480
	GCTGCTCTGG ACAGCACTGG GAGCGTGAGG CTGCTGTGCG CTTCCAGCAT GTGGGCACCT	540
30	CTGTGTTCCT GTCAAGTCACG GGTGAGCACT ATGGAAGCCC CATCCGTGGG CAGCATGAGG	600

166

TCCACGGCAT	GCCCAGTGCC	AACACGCACA	ATACGTGGAA	GGCCATGGAA	GGCATCTTCA	660
TCAAGCCTAG	TGTGGAGCCC	TCTGCAGGTC	ACCGATGAAC	CTGAGTGTGT	GGATGGATGG	720
GTGGATGGAG	GGTGGCAGGT	GGGGCGTCTG	CAGGGCCACT	CTTGGCAGAG	ACTTTGGGTT	780
TGTAGGGGTC	CTCAAGTGCC	TTTGTGATTA	AAGAATGTTG	GTCTATGAAA	AAAAAAAAAA	840
5	AAAAA					844

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 776 base pairs

10 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTCGAAATAA	AAGATCTGCT	CAAGAGAGCC	GCAGAAAAAG	AAGGTGTATG	TTGGGGGTTT	60
15	AGAGAGCAGG	GTCTTGAAAT	ACACAGCCCA	GAATATGGAG	CTTCAGAACAAAGTACAGCT	120
	TCTGGAGGAA	CAGAATTGT	CCCTCTAGA	TCAAATGAGG	AAACTCCAGG	180
	TGAGATATCA	AACAAAACCA	GCAGCAGCAG	CACCTGCATC	TTGGTCCTAC	240
	CTGCCCTCCTC	CTTGTACCTG	CTATGTACTC	CTCTGACACA	AGGGGGAGCC	300
	GCATGGAGTG	TTGTCCCGCC	AGCTTCGTGC	CCTCCCCAGT	GAGGACCCCT	360
20	GCTGCCCTGCC	CTGCAGTCAG	AAGTGGCGAA	AGACAGCACA	CACCACTGGT	420
	AGACTGTGTA	CTCCAGGCC	CTGGCAACAC	TTCCCTGCTG	CTGCATTACA	480
	TCCCAGTGCA	GAGCCTCCCC	TGGAGTGGCC	ATTCCTGAC	CTCTTCTCAG	540
	CCGAGGTCCC	ATCCTCCCCC	TGCAGGCAA	TCTCACAAGG	AAGGGAGGGAT	600
25	TGGTAGCCCC	TCTGTCAATT	TGCAGGACAG	ATACTCAGGC	TAGATATGAG	660
	GGGTCTCAGC	AGGAGCCTGG	GGGGCTCCCC	ATCTGTGTCC	AAATAAAAAG	720
	GGGCTGGCCG	CAGCTCCTGT	GCCCTGTCAG	GACGACTGAG	GGCTCAAACA	776

## (2) INFORMATION FOR SEQ ID NO: 14:

## 30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1376 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GAATTGGCGCA	CGAGGCGCCT	ACCCCTGCCTG	CAGGTGAGCA	GTGGTGTGTG	AGAGCCAGGC	60
GTCCCCCTCTGC	CTGCCCACTC	AGTGGCAACA	CCCGGGAGCT	GTTTTGTCTT	TTGTGGAGCC	120
TCAGCAGGTTTC	CCTCTTTCAAG	AACTCACTGC	CAAGAGCCCT	GAACAGGAGC	CACCATGCAG	180
TGCTTCAGCT	TCATTAAGAC	CATGATGATC	CTCTTCAATT	TGCTCATCTT	TCTGTGTGGT	240
10 GCAGCCCTGT	TGGCAGTGGG	CATCTGGGTG	TCAATCGATG	GGGCATCCTT	TCTGAAGATC	300
TTCGGGCCAC	TGTCGTCCAG	TGCCATGCAG	TTTGTCAACG	TGGGCTACTT	CCTCATCGCA	360
GCCGGCGTTG	TGGTCTTTC	TCTTGGTTTC	CTGGGCTGCT	ATGGTGTCAA	GACTGAGAGC	420
AAGTGTGCC	TCGTGACGTT	CTTCTTCATC	CTCCTCCTCA	TCTTCATTGC	TGAGGTTGCA	480
GCTGCTGTGG	TGCCCTGGT	GTACACCACA	ATGGCTGAGC	ACTTCCTGAC	GTTGCTGGTA	540
15 GTGCCTGCCA	TCAAGAAAAGA	TTATGGTTCC	CAGGAAGACT	TCACTCAAGT	GTGGAACACC	600
ACCATGAAAAG	GGCTCAAGTG	CTGIGGCTTC	ACCAACTATA	CGGATTTTGA	GGACTCACCC	660
TACTTCAAAG	AGAACAGTGC	CTTTCCCCCA	TTCTGTTGCA	ATGACAACGT	CACCAACACA	720
GCCAATGAAA	CCTGCCACCAA	CCAAAAGGCT	CACGACCAAA	AAGTAGAGGG	TTGCTTCAAT	780
CAGCTTTTGT	ATGACATCCG	AACTAATGCA	GTCACCGTGG	GTGGTGTGGC	AGCTGGAATT	840
20 GGGGGCCTCG	AGCTGGCTGC	CATGATTGTG	TCCATGTATC	TGTACTGCAA	TCTACAATAA	900
GTCCACTTCT	GCCTCTGCCA	CTACTGCTGC	CACATGGAA	CTGTGAAGAG	GCACCCCTGGC	960
AAGCAGCAGT	GATTGGGGGA	GGGGACAGGA	TCTAACAAATG	TCACCTGGGC	CAGAATGGAC	1020
CTGCCCTTTC	TGCTCCAGAC	TTGGGGCTAG	ATAGGGACCA	CTCCCTTTAN	GGCATGCCCTG	1080
ACTTTCTTC	CATTGGTGGG	TGGATGGGTG	GGGGCATTTC	CAGAGCCTCT	AAGGTAGCCA	1140
25 GTTCTGTTGC	CCATTCCCCC	AGTCTATTAA	ACCCCTGATA	TGCCCTCTAG	GCCTAGTGGT	1200
GATCCCAGTG	CTCTACTGGG	GGATGAGAGA	AAGGCATTTT	ATAGCCTGGG	CATAAGTGAA	1260
ATCAGCAGAG	CCTCTGGGTG	GATGTGTAGA	AGGCACCTCA	AAATGCATAA	ACCTGTTACA	1320
ATGTTAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAYTCG	AGGGGGGTCC	CGTACC	1376

## 30 (2) INFORMATION FOR SEQ ID NO: 15:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 502 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

5 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TAAAAACAGTG	CCTGCCCTCAA	AGGGAGGACT	CAGTCATAT	CTGTTGAATG	AATGAATGAA	60
TAATTGCGTG	GGTCAACGAA	TGAATGGCTG	AATGAATGAT	TTCTCCCTTC	CCTCGGCACT	120
GTCTGGAGTC	CCCAGGACAG	GCATGGGCAG	CAGTCGCTGG	TCTGTGGCCT	GTCCCCACTGG	180
10 ACTTGGGGTT	CTCATGCTTG	GTCTGGGCGG	AGATCACCCA	CCAGGCTCCC	AGGTCGATCC	240
TCTGCTCATG	GGAARCTGCG	TCCGGCCCN	GCTGCCAGAA	CTCACTGCAS	GGTGGAGGGA	300
ARARCAGGRA	CGATCTGCGA	GCCCTGAAAC	AGCGCACAAG	AGCCGAGGAG	CCGCTGCTTA	360
AAATGCAGGC	GTTGAGAGGA	GTTCGCGCTC	CTTTTTTGAG	TTGAATATGA	GATTTCCGAG	420
CAGCCATGAC	GAGTTGGGTT	GGTGGAAATG	GGGAGTCCGT	TCCTCAGTC	GATGGAGGAG	480
15 GGGGTCCCCT	TGGATCTCCT	CT				502

## (2) INFORMATION FOR SEQ ID NO: 16:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 425 base pairs

20 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ATCTCTAGTG	GTGGCTGCCG	TGCTCCAGA	CAATCGGAAT	CCTGCCCTCA	CCACCATGGG	60
25 CTGGCTTTTT	CTAAAGGTTT	TGTTGGCGGG	AGTGAGTTTC	TCAGGATTTC	TTTATCCTCT	120
TGTGGATTTT	TGCATCAGTG	GGAAAACAAG	AGGACAGAAC	CCAAACTTTC	TGATTATTTT	180
GGCCGATGAC	ATGGGGTGGG	GTGACTGGGG	AGCAAACTGG	GCAGAAACAA	AGGACACTGC	240
CAACCTTGAT	AAGATGGCTT	CGGAGGAAAT	GARGTGARTC	TTGARATGCC	ARGCCAGCTT	300
TCTTTGGAWG	TCTTACTCC	GTTCTTGAAA	AGGGAAAGGG	CCGTGCAAAG	CACTTAARGA	360
30 WTCATKGATG	GACCCATGTG	ATTTARTTAA	TTTATTAATT	AATTIGGTTT	GGAARCCAGC	420

ATAGC

425

## (2) INFORMATION FOR SEQ ID NO: 17:

## (i) SEQUENCE CHARACTERISTICS:

- 5                   (A) LENGTH: 1316 base pairs  
                   (B) TYPE: nucleic acid  
                   (C) STRANDEDNESS: double  
                   (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

10	GGCACGAGGA GCTGGGGGAG CCTGAGGTGC GCTACGTGGC TGGCATGCAT GGGAACGAGG	60
	CCCTGGGGCG GGAGTTGCTT CTGCTCCTGA TGCAGTTCT GTGCCATGAG TTCCCTGCGAG	120
	GGAACCCACG GGTGACCCGG CTGCTCTCTG AGATGCCAT TCACCTGCTG CCCTCCATGA	180
	ACCCTGATGG CTATGAGATC GCCTACCACC GGGGTTCAAGA GCTGGTGGGC TGGGCCGAGG	240
	GCCGCTGGAA CAACCAAGAGC ATCGATCTTA ACCATAATTT TGCTGACCTC AACACACAC	300
15	TGTGGGAAGC ACAGGACGAT GGGAAAGGTGC CCCACATCGT CCCCCAACCAT CACCTGCCAT	360
	TGCCCACTTA CTACACCCCTG CCCAATGCCA CGTGCGCTCC TGAAACGCGG GCAGTAATCA	420
	AGTGGATGAA GCGGATCCCC TTTGTGCTAA GTGCCAACCT CCACGGGGGT GAGCTCGTGG	480
	TGTCCTACCC ATTGACATG ACTCGCACCC CGTGGCGCTGC CGCGGAGCTC ACGCCCACAC	540
	CAGATGATGC TGTGTTTCGC TGGCTCAGCA CTGCTATGC TGGCAGTAAT CTGGCCATGC	600
20	AGGACACCAAG CGCCGACACC TGCCACAGCC AGGACTTCCTC CGTGCACGGC AACATCATCA	660
	ACGGGGCTGA CTGGCACACG GTCCCCGGGA GCATGAATGA CTTCAGCTAC CTACACACCA	720
	ACTGCTTTGA GGTCACTGTG GAGCTGTCT GTGACAAGTT CCCTCACGAG AATGAATTGC	780
	CCCAGGAGTG GGAGAACAAAC AAAGACGCC TCCTCACCTA CCTGGAGCAG GTGCGCATGG	840
	GCATTGCAGG ACTGGTGAGG GACAAGGACA CGGAGCTTGG GATTGCTGAC GCTGTCATTG	900
25	CCGTGGATGG GATTAACCAT GACGTGACCA CGGCGTGGGG CGGGGATTAT TGGCGTCTGC	960
	TGACCCCAAG GGACTACATG GTGACTGCCA GTGCCGAGGG CTACCAATTCA GTGACACGGA	1020
	ACTGTCGGGT CACCTTGAA GAGGGCCCT TCCCTGCAA TTTCGTGCTC ACCAAGACTC	1080
	CCAAACAGAG GCTGCGCGAG CTGCTGGCAG CTGGGGCAA GGTGCCCCCG GACCTTCGCA	1140
	GGCGCTTGGA GCGGCTAAGG GGACAGAAGG ATTGATACCT GCGGTTTAAG AGCCCTAGGG	1200
30	CAGGCTGGAC CTGTCAAGAC GGGAAAGGGGA AGAGTAGAGA CGGAGGGACA AAGTGAGGAA	1260

170

AAGGTGCTCA TTAAAGCTAC CGGGCACCTT AAAAAAAAAA AAAAAAAAAA AAAAAA 1316

## (2) INFORMATION FOR SEQ ID NO: 18:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 436 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

10	AAAAAAATTTC AATGGATATT ATGAAAATAA GAGAGTATT CCAGAAGTAT GGATATAGTC	60
	CACGTGCAA GAAAATTCA GTACACGAGC AAGAAGCCAT TAACTCTGAC CCAGAGTTGT	120
	CTAATTGTGA AAATTTTCAG AAGACTGATG TGAAAGATGA TCTGTCTGAT CCTCCTGTTG	180
	CAAGCAGTTG TATTCTGAG AAGTCTCCAC GTAGTCCACA ACTTTCAGAT TTTGGACTTG	240
	AGCGGTACAT CGTATCCCAA GTTCTACCAA ACCCTCCACA GGCAGTGAAC AACTATAAGG	300
15	AAGAGCCGT AATTGTAACC CCACCTACCA AACAACTCACT AGTAAAAGTA CTAAAAACTC	360
	CAAAATGTGC ACTAAAATGG ATGATTTGA GTGTGTACTC CTAAATTAGA ACACTTGGT	420
	ATCTCTGAAT ATACTA	436

## (2) INFORMATION FOR SEQ ID NO: 19:

## 20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 503 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

	TGTGCATATC CTGGGGAAAA AAATGGTACA TGTTTAGAA ATTTTACTGT TTATAACAAT	60
	GCAGGCAGTC AGTTTCCCGT TTCAAACACA GATAGATACA TCCAACACTC AAGATCTGCC	120
	AGAGAGGCAG CCAGCATCTA TTGTTAAAA AGGTTCAAA AAGAATTGG ATTGCTCKTT	180
	TCTCTTTGA ATCTGTGTGC CAAATGACAG GGACCAATAT TCGTCTTCCTT TTTCKGTAAG	240
30	AYTCAGAAAG AMACATGAAA GAACCCAGAA TGCATTTCTT AAAGGGATTT AGTGCAGTTA	300

TTTTAAATAA TTTATGCACG CACACACACA TACATATATC CCCCGAGTAC ATATTTTTTC	360
CCTTTTACT TGTGTGCAAT CAGTAGCTAC AATGACTGAA ATCCACTTCT TTGGGACTGT	420
GACATTAAAG CAAATCTTGT NTCTAGAAAN CGAAATGCCA NANTCTCGCA CAAAGCTGCT	480
CCGTCTGGGG CAACAAATCC ACA	503

5

## (2) INFORMATION FOR SEQ ID NO: 20:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 358 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GGGCTGTCTC CCCAGTAGTA ACTTGCTGGC CCTGCCCTG AAGTGGGAA ACTGTGAAGG	60
GCTCCTTGAT CAAGCTTGTC CTCTTTCTT ACCTCTTCCT CTCTTCTGTT TCCGCTGCAG	120
15 CTGAACAGGC CAGCAGGCAA CCTGCCATGG GGTCCTGTC CAAGAACCGG TCCTTCTTCT	180
GGATGACTGG GCTCCTGGTA TTCATCAGCC TCCTCCTCAG TGAGTGGCAG GGTCCCTGG	240
AAGGGAGGGC AATTGGAGAG GGCTGGCTA GCTGGCTCT GACCAACGGG TGGGCTGTT	300
AACTTCTGAT GTCTTTGGGC AACAAACACAG AAAAACACTC TGTTATGATT TACGAAAN	358

## 20 (2) INFORMATION FOR SEQ ID NO: 21:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1926 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

AGTGAAGGGA GCTGGCCGTG CGACTGGGCT TCGGGCCCTG TGCCAGAGGA GCANGCCTTC	60
CTGAGCAGGA GGAAGCAGGT GGTGGCCGCG GCCTTGAGGC AGGCCCTGCA GCTGGATGGA	120
30 GACCTGCAGG AGGATGAGAT CCCAGTAGTA GCTATTATGG CCACCTGGTGG TGGGATCCGG	180
GCAATGACTT CCCCTGTATGG GCAGCTGGCT GGCCTGAAGG AGCTGGGCCT CTGGATGTC	240

	KTCTCCTACA TCACCGGGGC CTCGGGCTCC ACCTGGGCCT TGGCCAACCT TTATAAGGAC	300
	CCAGAGTGGT CTCAGAAAGGA CCTGGCAGGG CCCACTGAGT TGCTGAAGAC CCAGGTGACC	360
	AAGAACAAAGC TGGGTGTGCT GGCCCCCAGC CAGCTGCAGC GGTACCGGCA GGAGCTGGCC	420
	GAGCGTGCCTC GCTTGGGCTA CCCAAGCTGC TTCACCAACC TGTTGGCCCT CATCAACGAG	480
5	GCGCTGCTGC ATGATGAGCC CCATGATCAC AAGCTCTCAG ATCAACGGGA GGCCCTGAGT	540
	CATGGCCAGA ACCCTCTGCC CATCTACTGT GCCCTCAACA CCAAAGGGCA GAGCCTGACC	600
	ACTTTTGAAAT TTGGGGAGTG GTGCGAGTTC TCTCCCTACG AGGTGGCTT CCCCAAGTAC	660
	GGGGCCTTCA TCCCCCTCTGA CCTCTTTGGC TCCGAGTTCT TTATGGGCA GCTGATGAAG	720
	AGGCTTCCTG AGTCCCCTCAT CTGCTTCTTA GAAGGTATCT GGAGCAACCT GTATGCAGCC	780
10	AACCTCCAGG ACAGCTTATA CTGGGCTTCA GAGCCCAGCC AGTTCTGGGA CGCGTGGGTC	840
	AGGAACCAGG CCAACCTGGA CAAGGAGCAG GTCCCCCTTC TGAAGATAGA AGAACCAACCC	900
	TCAACAGCCG GCAGAAATAGC TGAGTMTTTC ACCGATCTTC TGACGTGGCG TCCACTGGCC	960
	CAGGCCACAC ATAATTTCCT GCGTGGCCTC CATTCCACA AAGACTACTT TCAGCATCCT	1020
	CACTTCTCCA CATGGAAAGC TACCACTCTG GATGGGCTCC CCAACCAGCT GACACCTCG	1080
15	GAGCCCCACC TGTGCCTGCT GGATGTTGGC TACCTCATCA ATACCAGCTG CCTGCCCTC	1140
	CTGCGACCCA CTCGGACGT GGACCTCATC CTGTCATTGG ACTACAACCT CCACGGAGCC	1200
	TTCCAGCAGT TGCAGCTCCT GGGCCGGTTC TGCCAGGAGC AGGGGATCCC GTTCCCACCC	1260
	ATCTCGCCCA GCCCCGAAGA GCAGCTCCAG CCTCGGGAGT GCCACACCTT CTCCGACCCC	1320
	ACCTGCCCG GAGCCCTGC GGTGCTGCAC TTTCTCTGG TCAGCGACTC CTTCCGGGAG	1380
20	TACTCGGCCC CTGGGTCCG GCGGACACCC GAGGAGGGGG CAGCTGGGA CGTGAACCTG	1440
	TCTTCATCGG ACTCTCCCTA CCACTACACG AAGGTGACCT ACAGCCAGGA GGACGTGGAC	1500
	AAGCTGCTGC ACCTGACACA TTACAATGTC TGCAACAACC AGGAGCAGCT GCTGGAGGCT	1560
	CTGCGCCAGG CAGTGCAGCG GAGGCCAG CGCAGGCCCC ACTGATGGCC GGGGCCCTG	1620
	CCACCCCTAA CTCTCAITCA TTCCCTGGCT GCTGAGTTGC AGGTGGAAC TGTCATCAGC	1680
25	CAGTGCTTNC AGAGCCTCGG GCTCAGGTGG CACTGTCCCA GGGTCCAGGC TGAGGGCTGG	1740
	GAGCTCCCTT GCGCCTCAGC AGTTTGCACT GGGGTAAGGA GGCCAAGCCC ATTITGTTAA	1800
	TCACCCAAA CCCCCCGGCC TGTGCCTGTT TTCCCTCTG CGCTACCTTG AGTAGTTGGA	1860
30	GCACTTGATA CATCACAGAC TCATACAAAT GTGAGGGCT GAGAAAAAAA AAAAAAAA	1920
	ACTCGA	1926

## (2) INFORMATION FOR SEQ ID NO: 22:

5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1224 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CGGCCGAAGC TCCGTCGGCG CCGCGGCCGG CTCCGCTCA CCTCCCCGCC GCGGCTGCC	60
TCTGCCCGG TTGTCCAAGA TGGAGGGCGC TCCACCGGGG TCGCTCGCCC TCCGGCTCCT	120
GCTGTTCTG GCGCTACCCG CCTCCGGCTG GCTGACGACG GGCGCCCCCG AGCCGCCGCC	180
GCTGTCCGGA GCCCCACAGG ACGGCATCAG AATTAATGTA ACTACACTGA AAGATGATGG	240
GGACATATCT AACACAGCAGG TTGTTCTTAA CATAACCTAT GAGAGTGGAC AGGTGTATGT	300
AAATGACTTA CCTGTAAATA GTGGTGTAAAC CCGAATAAGC TGTCAGACTT TGATAGTGAA	360
GAATGAAAAT CTTGAAAATT TGGAGGAAAA AGAATATTTT GGAATTGTCA GTGTAAGGAT	420
TTTAGTTCAT GAGTGGCCTA TGACATCTGG TTCCAGTTTG CAACTAATTG TCATTCAAGA	480
AGAGGTAGTA GAGATTGATG GAAAACAAGT TCAGCAAAG GATGTCACTG AAATTGATAT	540
TTTAGTTAAG ACCGGGGAG TACTCAGACA TTCAAACAT ACCCTCCCTT TGGAAAGAAAG	600
CATGCTCTAC TCTATTTCTC GAGACAGTGA CATTTTATTT ACCCTTCCTA ACCTCTCCAA	660
AAAAGAAAGT GTTAGTTCAC TGCAAACAC TAGCCAGTAT CTTATCAGGA ATGTGGAAAC	720
CACTGTAGAT GAAGATGTTT TACCTGGCA AGTTACCTGA AACTCCTCTC AGAGCAGAGC	780
CGCCATCTTC ATATAAGGTA ATGTGTCAGT GGATGGAAAA GTTTAGAAAA GATCTGTGTA	840
GGTTCTGGAG CAACGTTTTC CCAGTATTCT TTCAAGTTTT GAACATCATG GTGGTTGGAA	900
TTACAGGAGC AGCTGTGGTA ATAACCATCT TAAAGGTGTT TTTCCAGTT TCTGAATACA	960
AAGGAATTCT TCAGTTGGAT AAAGTGGACG TCATACCTGT GACAGCTATC AACTTATATC	1020
CAGATGGTCC AGAGAAAAGA GCTGAAAACC TTGAAGATAA AACATGTATT TAAAACGCCA	1080
TCTCATATCA TGGACTCCGA AGTAGCCTGT TGCCTCCAAA TTTGCCACTT GAATATAATT	1140
TTCTTTAAAT CGTTAAGAAT CAGTTTATAC ACTAGAGAAA TTGCTAAACT CTAAGACTGC	1200
CTGAAAATTG ACCTTTACAG TGCC	1224

55

## (2) INFORMATION FOR SEQ ID NO: 23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 694 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

5	GGCACGAGTC TTATTGTGCA CTGTAGCCTG AATCCCCAG GGTAAATTAAAT ATGAAGTGCA	60
	AAAAGTTGAA TGTTCAGTC TAAAAGGCAG TGGGAGAAAT TACATAGCAT GGAAATAATA	120
10	AAATGAACTC TTATTAATGA GAACGAGGCT CTTGCAGTGG CAAGTTCTGC TGGTCACCCG	180
	ATGGGGATGG GACCCTTCAG AGCTTTTTT TGGGTAATAC TCACAGTTTC CAACGTCIGT	240
15	GTACTTTCA AAATGAGCTT GTTCTTCCTT CTGACACTCA TCTCAAAGCT CCATGGTGAC	300
	GCAGAGGTCT GTTGAAGGTC ACAGGTCCCTC GCTTGCATTG GCATACGGTC CTGTAGCATC	360
20	ACTTGTAGC CCACTGCTGC TTGAAGGAAC TAAGAGTATT CAGGGATAGA GAGCTGAAAA	420
	TAGGATTAAT TCCTTCCTTT TGACTCTCCC CTCAAGATGT CCTTGCTTTG GTCTGAAAAC	480
25	CTCTCCGTAC AACCTTTGCC CAAAGCAAAC CATCTGCCIT TTCTGAACTC TGAGTGAAATA	540
	TATTAGCATC TTCCCTCTG AGCCCTCGTA CTGCCANGTT TGTTGTTTG TTTGTTCCA	600
30	AGAGACTGTG TCTTGCTCTG TCACCCAGGA GTTGAAACC AGCCTGGCAA CATAGCAAGA	660
	CCCTATCTCT ACAAAAAAAAA AAAAAAAAAA AAAA	694

35

## (2) INFORMATION FOR SEQ ID NO: 24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 796 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

40	ATGAGCGGCG GTTGGATGGC GCAGGTTGGA GCGTGGCGAA CAGGGGCTCT GGGCTGGCG	60
45	CTGCTGCTGC TGCTCGGCCT CGGACTAGGC CTGGAGGCCG CGCGAGCCCG CTTCACCC	120
	CGACCTCTGC CCAGGCCGCA CCCGAGCTCA GGCTCGTGCC CACCCACCAA GTTCCAGTGC	180
50	CGCACCACTG GCTPATGCGT GCCCCTCACC TGGCGCTGCG ACAGGACTTG GACTGCAGCG	240
	ATGGCAGCGA TGAGGAGGAG TGCAGGATTG AGCCATGTAC CCAGAAAGGG CAATGCCAC	300
55	CGCCCCCTGG CCTCCCCCTGC CCCTGCACCG GCGTCAGTGA CTGCTCTGGG GGAACGTACA	360
	AGAAAATGCG CAACTGCAGC CGCCTGGCCT GCCTAGCAGS GRAGSKCMCG WKGCACGCTG	420

	AGCGATGACT GCATTCAC TACGTGGCGC TGCGACGGCC ACCCAGACTG TCCCGACTCC	480
5	AGCGACGAGC TCGGCTGTGG AACCAATGAG ATCCTCCCGG AAGGGGATGC CACAACCATG	540
	GGGCCCCCTG TGACCCCTGGA GAGTGTCAAC TCTCTCAGGA ATGCCACAAC CATGGGGCCC	600
	CCTGTGACCC TGGAGAGTGT CCCCTCTGTC GGGAAATGCCA CATCCTCCTC TGCCGGAGAC	660
10	CAGTCTGGAA GCCCAACTGC CTATGGGGTT ATTGCAGCTG CTGCGGTGCT CAGTGCAAGC	720
	CTGGTCACCG CCACCCCTCCT CCTTTTGTCC TGGCTCCGAG CCCAGGAGCG CCTCCGCCA	780
	CTGGGGTTAC TGGTGG	796

15

## (2) INFORMATION FOR SEQ ID NO: 25:

20

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 662 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

30	TAATTCCGCA CGAGGCTGTG GTGGAGAAGG ACGTGGCGTG CCGCTGGGTT CTGAGCCGGA	60
	GTTGGTCGGTG GGTGGGATGG AGGCGACCTT GGAGCAGCAC TTGGAAGACA CAATGAAGAA	120
	TCCCTCCATT GTTGGAGTCC TGTGCACAGA TTCACAAGGA CTTAACCTGG GTTGGCCCGG	180
35	GACCTGTCA GATGAGCATG CTGGAGTGAT ATCTGTTCTA GCCCAGCAAG CAGCTAACGCT	240
	AACCTCTGAC CCCACTGATA TTCCCTGTGGT GTGTCTAGAA TCAGATAATG GGAACATTAT	300
	GATCCAGAAA CACCGATGGCA TCACGGTGCG AGTGCACAAA ATGGCCCTTT GATGCTCATA	360
40	TCTGTTCTTC AGCAGCCTGT CATAGGAACT GGATCCTACC TATGTTAATT ACCTTATAGA	420
	ACTACTAAAG TTCCAGTAGT TAGGCCATTC ATTTAACATG CATTAGGCAC TTTTCTGT	480
45	ATTTAACAGT CAATGCTTT CTAATGCTCT ATGGACCGAC TATCAAGATA TTAGTAAGAA	540
	AGGATCATGT TTTGAAGCAG CAGGTCCAGG TCACTTTGTA TATAGAATTG TGCTGTATTC	600
	AATAAAATCTG TTTGGAGGAA AAAAAAAAAA AAAAAAAATTA CTGCGGNCCG ACAAGGGAAT	660
50	TC	662

55

## (2) INFORMATION FOR SEQ ID NO: 26:

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1105 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

5	CCTGATCCTC TCTTTCTGC AGTTCAAGGG AAAGACGAGA TCTTGACCAA GGCACTCTGC	60
	TTCTGCCCTT GGCTGGGAA GGGTGGCATG GAGCCTCTCC GGCTGCTCAT CTTACTCTTT	120
10	GTCACAGAGC TGTCCGGAGC CCACAACACC ACAGTGTCC AGGGCGTGGC GGGCCAGTCC	180
	CTGCAGGTGT CTTGCCCTA TGACTCCATG AAGCACTGGG GGAGGCGCAA GGCCTGGTGC	240
15	CGCCAGCTGG GAGAGAAGGG CCCATGCCAG CGTGTGGTCA GCACGCACAA CTTGTGGCTG	300
	CTGTCCTTCC TGAGGAGGTG GAATGGGAGC ACAGCCATCA CAGACGATAAC CCTGGGTGGC	360
	ACTCTCACCA TTACGCTGCG GAATCTACAA CCCCATGATG CGGGCTCTA CCAGTGCCAG	420
20	AGCCTCCATG GCAGTGAGGC TGACACCCCTC AGGAAGGTCC TGGTGGAGGT GCTCGCAGAC	480
	CCCCTGGATC ACCGGGATGC TGGAGATCTC TGGTTCCCCG GGGAGTCTGA GAGCTTCGAG	540
	GATGCCATG TGGACACAG CATCTCCAGG AGCTCTTCKT AGGAAAGGCC GCAAATTCCC	600
25	ATTCCCTCCC CTCTTGCTA TCYTTCTCCT CCAAGAYCTG CATCTTCTC ATCAAGATTC	660
	TAGCAGCCAG CGCCCTCTGG GCTGCAGCCT GGCATGGACA GAAGCCAGGG ACACATCCAC	720
30	CCAGTGAACG GGACTGTGGC CATGACCCAG GGTATCAGCT CCAAACCTTG CCAGGGCTGA	780
	GAGACACCTG AAGGAAGATG ATGGGAGGAA AAGCCCAGGA GAAGTCCAC CAGGGACCAG	840
	CCCAGCCTGC ATACTTGCCA CTTGGCCACC AGGACTCCTT GTTCTGCTCT GGCAAGAGAC	900
35	TACTCTGCCT GAACACTGCT TCTCCTGGAC CCTGGAAGCA GGGACTGGTT GAGGGAGTGG	960
	GGAGGTGGTA AGAACACCTG ACAACTTCTG AATATTGGAC ATTTTAAACA CTTACAAATA	1020
40	AATCCAAGAC TGTCAATTTT AAAAAAAAAA AAAAAAAAAMA AAARRRRRC CCCGGTACCC	1080
	AATTGCCCCAT ATAGTGAGTC GTATA	1105

45

## (2) INFORMATION FOR SEQ ID NO: 27:

50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1017 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:	
	CTCGCTGGG CTGTTCCCCG CCTTCATTC TCCCGACTCA GCTTCCACC CTGGGCTTTC	60
60	CGAGGTGCTT TCGCCGCTGT CCCCACCACT GCAGCCATGA TCTCCTAAC GGACACGCAG	120

	AAAATTGGAA TGGGATTAAC AGGATTTGGA GTGTTTTCC TGTTCTTGG AATGATTCTC	180
	TTTTTTGACA AAGCACTACT GGCTATTGGA AATGTTTAT TTGTAGCCGG CTTGGCTTT	240
5	GTAATTGGTT TAGAAAGAAC ATTCAAGATTTC TTCTTCCAAA AACATAAAAT GAAAGCTACA	300
	GGTTTTTTTC TGGGTGGTGT ATTTGTAGTC CTTATTGGTT GCCCTTTGAT AGGCATGATC	360
10	TTCGAAATT ATGGATTTTT TCTCTTGITC AGGGGCTTCT TTCCCTGTCGT TGTTGGCTTT	420
	ATTAGAAGAG TGCCAGTCCT TGGATCCCTC CTAAATTTAC CTGGAATTAG ATCATTIGTA	480
	GATAAAGITG GAGAAAGCAA CAATATGGTA TAACAACAAG TGAATTGAA GACTCATTIA	540
15	AAATATITGTG TTATTTATAA AGTCATTTGA AGAATATTCA GCACAAAATT AAATTACATG	600
	AAATAGCTTG TAATGTTCTT TACAGGAGTT TAAAACGTAT AGCCTACAAA GTACCAGCAG	660
20	CAAATTAGCA AAGAAGCAGT GAAAACAGGC TTCTACTCAA GTGAACTAAG AAGAAGTCAG	720
	CAAGCAAACG GAGAGAGGTG AAATCCATGT TAATGATGCT TAAGAAACTC TTGAAGGCTA	780
	TTTGTGTTGT TTTTCCACAA TGTGCGAAC TCAGCCATCC TTAGAGAACT GTGGTGCCTG	840
25	TTTCTTTCT TTTTATTTTG AAGGCTCAGG AGCATCCATA GGCAATTGCT TTTTAGAAAT	900
	GTCCACTGCA ATGCCAAAAA TATTTCCAGT TGCACTGTAT CTCTGGAAGT GATGCATGAA	960
	TTCGATTGGA TTGTGTCATT TTAAAGTATT AAAACCAAGG GAAACCCCAA AAAAAAA	1017
30		

## (2) INFORMATION FOR SEQ ID NO: 28:

35

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 391 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

45	CCCTGGAAAG AGGAACGTGAT GTTTGAGGGG ACAGATGTGG GTCACTTTCC CTGGCAGTGC	60
	CCTCTAGCCT TGCTGCCCTG GCTTTCTGAC CCCTTCCAGG CTTCAGGGGC CTGGGAGATC	120
	TCATGCCCTCA GCCCAGGAAA CATTAAATAG GGAAAGCAGA GACATGTCAT GTCAGCCCCA	180
50	CAGACAAGAA TTTCTAGAGC ACTTGTCTTG TTGTTCCCTG CCCCGACATT ACTCAGTCTG	240
	GGCCATGGAA TCCATCCAAT AAACACAGCA ACACCCCTATG NTACTGACCA AGCAAAGCTT	300
	GCCCCCTGGTA CCAAAGAGCT AAATCATGAC CAAAGTGTGA CATGAATGTA ACTGAAATGC	360
55	GGGTTAGTTG CTCAATGTAT GCAAAGTCCC A	391

## (2) INFORMATION FOR SEQ ID NO: 29:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1139 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

10	GGTGATATCT TCATAGTGGG CTATTACAGG CAGGAAAATG TTTTAACTGG TTTACAAAAT	60
	CCATCAATAAC TTGTGTCATT CCCTGTAAAA GGCAGGAGAC ATGTGATTAT GATCAGGAAA	120
15	CTGCCACAAAA TTATTTGTTT CAGCCCCCGT GTTATTGTCC TTTTGAACCTG TTTTTTTTTT	180
	ATTAAAGCCA AATTTGTGTT GTATATATTC GTATTCCATG TGTTAGATGG AACCATTTCC	240
20	TATCCAGTGT GAATAAAAAG AACAGTTGTA GTAAATTATT ATAAGCCGA TGATATTTCA	300
	TGGCAGGGTA TTCTACCAAG CTGTGCTTGT TGGTTTTTCC CATGACTGTA TTGCTTTTAT	360
	AAATGTACAA ATAGTTACTG AAATGACGAG ACCCTTGTT GCACAGCATT AATAAGAAC	420
25	TTGATAAGAA CCATATTCTG TTGACAGCCA GCTCACAGTT TCTTGCTGTA AGCTTGGTGC	480
	ACCCTCCAGT GAGACACAAG ATCTCTCTTT TACCAAAGTT GAGAACAGAG CTGGTGGATT	540
30	AATTAATAGT CTTCGATATC TGGCCATGGG TAACCTCATT GTAACTATCA TCAGAATGGG	600
	CAGAGATGAT CTTGAAGTGT CACATACACT AAAGTCCAAA CACTATGTCA GATGGGGTA	660
	AAATCCATTA AAGAACAGGA AAAAATAATT ATAAGATGAT AAGCAAATGT TTCAGCCAA	720
35	TGTCAACCCA GTTAAAAAAA AAATTAATGC TGTGTAAAAT GGTTGAATTG TTITGCAAAC	780
	TATATAAAGA CATATGCAGT AAAAAGTCTG TTAATGCACA TCCTGTGGGA ATGGAGTGT	840
40	CTAACCAATT GCCTTTCTT GTTATCTGAG CTCTCCTATA TTATCATACT CAGATAACCA	900
	AATTTAAAGA ATTAGAATAT GATTMTTAAT ACACITTAACA TTAAACTCTT CTAACTTCT	960
	TCTTTCTGTG ATAATTCAAG AGATAGTTAT GGATCTTCAA TGCTCTGAG TCATTGTTAT	1020
45	AAAAAATCG TTATCACTAT ACCATGCTAT AGGAGACTGG GCAAAACCTG TACAATGACA	1080
	ACCCCTGGAAG TTGCTTTTTT TAAAAAAATA ATAAATTTCT TAAATCAAAA AAAAAAAA	1139

50

## (2) INFORMATION FOR SEQ ID NO: 30:

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 465 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

	CCACGGGTCC GCGGACGCGT GGGGAAGGTT TGTGCCAGTA GACATTATGT TACTAAATCA	60
5	GCACTTTAAA ATCTTGGTT CTCTAATTCA TATGAATTG CTGTTGCTC TAATTTCTTT	120
	GGGCTCTCT AATTGAGTG GAGTACAATT TTGTTGTGAA ACAGTCCAGT GAAACTGTGC	180
	ACGGAAATGA AGGTAGAATT TTGGGAGGTA ATAATGATGT GAAACATAAA GATTTAATAA	240
10	TTACTGTCCA ACACAGTGGA GCAGCTTGTC CACAAATATA GTAATTACTA TTTATTGCTC	300
	TAAGGAAGAT TAAAAAAAGA TAGGGAAAAG GGGGAAACTT CTTTGAAAAA TGAAACATCT	360
15	GTTACATTA TGTCTAATTA TAAAATTTA ATCCTTACTG CATTCTTCT GTTCCTACAA	420
	ATGTATTAAA CATTCAAGTTT AACTGGTAAA AAAAAAAAAA AAAAAA	465

20

## (2) INFORMATION FOR SEQ ID NO: 31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

30	GCAACAAGCG GCCCACCTTC CTGAACATCA AGAACGCCACT GTCGTACCGC AAGCCCATGG	60
	ACACGGACCT GGTGTACATC GAGAAGTCGC CCAAACACTTG CGAGGAGGAC CCGGTGACCG	120
35	GCAGTGTGGG CACCCAGGGC CGCCGCTGCA ACAAGACGGC TCCCCAGGCC AGGGGCTGTG	180
	ACCTCATGTG CTGTGGCGT GGCTACAACA CCCACCAGTA CGCCCGCGTG TGGCAGTGCA	240
40	ACTGTAAAGTT CCACCTGGTC TGCTATGTCA AGTGCAACAC GTGCAGCGAG CGCACGGANG	300
	ATGTACACGT GCAAGTGAGC CCCGTGTGCA CACCACCCCTC CCGCTGCAAG TCAGATTGCT	360
	GGGAGGACTG GACCGTTTCC AAGCTGGGG CTCCCTGGCA GGATGCTGAG CTTGTCTTTT	420
45	CTGCTGAGGA GGGTACTTTT CCTGGTTTC CTGCAGGCAT CCGTGGGGGA AAAAAAATCT	480
	CTCAGAGNCC TCAACTATTC TGTCCACAC CCAATGCTGS TCCACCCCTCC CCCAGACACA	540
	GCCCAGGTCC CTCCGCGGCT GGAGCGAAGC CTTCTGCAGC AGGAACCTCTG GACCCCTGGG	600
50	CCTCATCACA GCAATATTTA ACAATTATT CCTGATAAAA ATAATATTAAT TTTATTTAAT	660
	AAAAAAGAAT TCTTCCAAAA AAAAAAAAAA AAAAAAACNT CG	702

55

## (2) INFORMATION FOR SEQ ID NO: 32:

60 (i) SEQUENCE CHARACTERISTICS:

180

- (A) LENGTH: 1142 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CGGCACGAGG AAGAAATGGC AGAGACTGGA ATCTCTCTTC ATGAAAAAT GCAGCCCTT	60
10 AACATTCAGTT CGACAGAGTG CAGCTCCCTTC TCTCCACCCA CCACAGTGAT TCTCCTTATC	120
CTGCTGTGCT TTGAGGGCCT GCTCTTCCTC ATTTTCACAT CAGTGATGTT TGGGACCCAG	180
15 GTGCACTCCA TCTGCACAGA TGAGACGGGA ATAGAACAAAT TGAAAAGGA AGAGAGAAGA	240
TGGGCTAAAA AAACAAAATG GATGAACATG AAAGCCGTTT TTGGCCACCC CTTCTCTCTA	300
GGCTGGGCCA GCCCCTTGC CACGCCAGAC CAAGGGAGG CAGACCCGTA CCAGTATGTG	360
20 GTCTGAAGGA CCCCGACCGG CATGCCACT CAGACACAAG TCCACACCAC AGCACTACCG	420
TCCCCATCCGT TCTCATGAAT GTTTAAATCG AAAAACAAA ACAACTACTC TTAAAACCTT	480
25 TTTTATGTCT CAAGTAAAAT GGCTGAGCAT TGCAGAGARA AAAAAAAGTC CCCACATTTC	540
ATTTTTTAAA ACCATCCTT TCGATTTCTT TTGGTGACCG AAGCTGCTCT CTTTTCCCTT	600
TAAAATCACT TCTCTGGCCT CTGGTTCTC TCTGCTGTCT GTCTGGCATG ACTAATGTAG	660
30 AGGGCGCTGT CTCGGCTGTG GCCCATTCTA CTAACTGAGT GAGACATGAC GCTGTGCTGG	720
GATGGAATAG TCTGGACACC TGGTGGGGGA TCCATGGAA AGCCAGGAGG GCCCTGACCT	780
35 TCCCCACTGCC CAGGAGGCAG TGGCGGGCTC CCCGATGGGA CATAAAACCT CACCGAAGAT	840
GGATGCTTAC CCCCTTGAGGC CTGAGAAAGGG CAGGATCAGA AGGGACCTTG GCACAGCGAC	900
CTCATCCCCC AAGTGGACAC GGTGTTGCTG CTAACTCGCA AAGCAATTGC CTGCTTGTA	960
40 CTTTATGGG TTGGGGTGTG TAGAATGATT TTGGGGGGGA GTGGGGGAGA AAGATGAAAG	1020
AGGTCTTATT TGTATTCTGA ATCAGCAATT ATATTCCTG TGATTATTG GAAGAGTGTG	1080
45 TAGGAAAGAC GTTTTTCCAG TTCAAAATGC CTTATACAAT CAAGAGGAAA AAAAAAAAAA	1140
AG	1142

50

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 928 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

	GGCACGAGGT CTAATGAGGG CTCTCTTGT TGCTAGAGAT GAGAGAAATG TATACTAAC	60
	ATTTTAATTT GTACTTAAAAA TACATTTTAC TAATCATATT GATTTTAAAT ATGACAAATT	120
5	CTTCTAGTAG ATACTAATCT TTCTTGTAA TCATATTGTC CTAGAGAAC CTAGGTAAAA	180
	ATGGGTTCCA CCTAGTCTGT TTGTATAACA CCTTCCCCCG TCCCTCTCC ATCCCTGCCA	240
10	ATTGGGCTCT ATGCATATTG ACAAGCAAAT AAGAAAACCT TAGGTTCTTG TATTGAATT	300
	TCCAAAACAA TAAAAGGTTT TGACTCAAGA TTTGCATTCA AGAAGAGGCA GAAATTTGT	360
	CTTATCTTTT TATCATTGG TGAACTTGTG TTTCTCTGTA TGCTTAGAAA ATTTACACAC	420
15	AAGGAATGTT TGAAAAAGTG AGAATTTAG AGTGCTTGGG TGGTTTTAT TTGGTCAGTG	480
	CTGATGTGTT AGGTGTTAG GGAAATAATG CTTCAGGACC TTTTGACAA CACAGCTICA	540
20	TGAATGACTG GGGATATTT ATGTTTGTGC TGAGAAAAGG GAGGGAGTGG GCAGGTTGGA	600
	GTGGGGACCT TTCCATTGAA AGCAGTGCAG TCAGCTGTT CGTAGATGCA TTTTTCTTT	660
	ATGCTTGTAA CATTGTTCTT GTGTCCATAA TTGACTGAA TGTCAGCTC CAGGAATGCA	720
25	AGGCATTTAT CAGGTGACCA GAAGTAGAAC CTTGTTGATT ATGAAATGGA AGAATAATGT	780
	CAAGGTAGTG GGGTAAAAT GACAAATAAG ATTTTACTGG TGAATTCTCA TGCTTAGTAT	840
	GTACATTAAC CTCTTTTAA GTTGCATGTT AATCTGGTAT AACGTATTGT GTCTGGTTA	900
30	TGCTTTGAGT AAAAAAAAAA AAAAAAAA	928

35

(2) INFORMATION FOR SEQ ID NO: 34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 773 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

45	GGCACGAGGT CTGGCCTCTC ATTTCCCTAC ACTCTGACAT GAATGAATTA TTATTATTTT	60
	TCTTTTTCTT TTTTTTTTTT ACATTTGTGTA TAGAAACAAA TTCATTTAAA CAAACTTATT	120
50	ATTATTATTT TTTACAAAAT ATATATATGG AGATGCTCCC TCCCCCTGTG AACCCCCAG	180
	TGCCCCCGTG GGGCTGAGTC TGTGGGCCAAG CTGGATTCTG TGTACCTAGT	240
	ACACAGGCAT GACTGGGATC CCGTGTACCG AGTACACGAC CCAGGTATGT ACCAAGTAGG	300
55	CACCCCTGGG CGCACCCACT GGGGCCAGGG GTCGGGGAT GTTGGAGGCC TCCTCCCCAC	360
	CCCACCTCCC TCACTTCACT GCATTCCAGA TTGGACATGT TCCATAGCCT TGCTGGGAA	420
60	GGGCCCCACTG CCAACTCCCT CTGCCCCAGC CCCACCCCTG GCCATCTCCC TTTGGGAACT	480

	AGGGGGCTGC TGGTGGAAA TGGGAGCCAG GGCAGATGTA TGCATTCCCTT TATGTCCCTG	540
5	TAAATGTGGG ACTACAAGAA GAGGAGCTGC CTGAGTGGTA CTTTCTCTTC CTGGTAATCC	600
	TCTGGCCCCAG CCTTATGGCA GAATAGAGGT ATTTTTAGGC TATTTTTGTA ATATGGCTTC	660
	TGGTCAAAAT CCCTGTGTAG CTGAATTCCC AAGCCCTGCA TTGTACAGCC CCCCCACTCCC	720
10	CTCACCACT AATAAAGGAA TAGTTAACAC TCAAAAAAAA AAAAAAAA AAA	773

## 15 (2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 453 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

25	TAAAATGTTA CACGCTTGTC ATATTCCAGG CACTGCACTA TGTATGCCGT TTATCAACAG	60
	TTAGCTCAGC TAACCCTCAT GGTAAACCTTG TTAGCCCCGA TTTTGCCAGA TGAGCAAAGT	120
30	GAGGTTTTG AGGCCTTAAG TAACTTGCCC AAGGTACAGT GGCTGGGAAG TAACTCTCCC	180
	AGTTCTGAGA TGCCCGAGCC TGGACGCTTT GTCATTGTAC ACCATCAACT CAGTGCTGCC	240
	AGTCATTCCA GCAGCCAGCT AGCGTAGTCA AGGTTTCTCC ACCTTAGCAC TGTGACATT	300
35	TCGAGCCAGA TAATTCTCTG TGGTGAGGAG CTGTCCTATG CCTTGTAGGA TATACAACAG	360
	CATCYTGGCT TTACCCACCA GATGYTGGAA CACCTCCCCA GTCGTGACAG CCCAAAATGT	420
	CTATAGACGT TGCCACGTAT ACCCAGGGGT TCC	453
40		

## (2) INFORMATION FOR SEQ ID NO: 36:

- 45 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 459 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

55	GTGACTGCCG CCCTGCCCGC AGCCATGTGG CCCCCGCTGT TGCTGCTGCT GCTGCTGCTC	60
	CGGGCCGCC CGGTCCCCAC CGCCAAAGCC GCTCCCCACC CGGATGCTAA CACCCAGGAA	120
	GGCCTTCAGA ACCTGCTCCA AGGAGTCGGG GCTGGCGGAG ACGGAGAGCT GCGGGCAGAC	180
60	TCACACCTGG CCCCGGGCTC TGGCTGTATT GATGGGGCTG TGGTGGCCAC GCGACCAGAA	240

	AGCCGGGGAG GAAGACCTGC GGTTCCGTGA GAGGCGTCCA GGGCTGCAGG CCACGGCGAC	300
5	AGGCTCCGGG GAACATGGGG CTTTCCCTGT CCACTCCAA GGAGTGTGGG CCTCAACGCA	360
	TTGGCAGGGG ACGGCCGTGT GCCCTCTYCA GACCCCACCC CCAGATGCAT TTATTAGAAA	420
	TAATAAATTC TTTCTTAGCT AAAAAAAAAA AAAAAAAAT	459

10

## (2) INFORMATION FOR SEQ ID NO: 37:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 509 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

	ATGAAATTAA CCACTCTCCT CTTCTTGGCA GCTGTAGCAG GGGCCCTGGT CTATGCTGAA	60
25	GATGCCTCCT CTGACTCGAC GGGTGCTGAT CCTGCCAGG AAGCTGGAC CTCTAACGCCT	120
	AATGAAGAGA TCTCAGGTCC AGCAGAACCA GCTTCACCCCC CAGAGACAAC CACAACAGCC	180
	CAGGAGACTT CGGGGGCAGC AGTTCAAGGG ACAGCCAAGG TCACCTCAAG CAGGCAGGAA	240
30	CTAAACCCCC TGAAATCCAT AGTGGAGAAA ACTATCTTAC TAACAGAACCA AGCCCTTGCA	300
	AAAGCAGGAA AAGGAATGCA CGGAGGCGTG CCAGGTGGAA AACAAATTCAAT CGAAAATGGA	360
35	AGTGAATTTC CACAAAAATT ACTGAAGAAA TTCAGTCTAT TAAAACCATG GGCAATGAGAA	420
	GCTGAAAAGA ATGGGATCAT TGGACTTAAA GCCTTAAATA CCCTTGTAGC CCAGAGCTAT	480
40	TAAAACGAAA GCATCCAAAA AAAAAAAAAA	509

## (2) INFORMATION FOR SEQ ID NO: 38:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 598 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

	ATGTTGGGCT GTGGGATCCC AGCGCTGGGC CTGCTCCTGC TGCTGCAGGG CTCGGCAGAC	60
55	GGAAATGGAA TCCAGGGATT CTTCTACCCA TGGAGCTGTG AGGGTGACAT ATGGGACCGG	120
	GAGAGCTGTG GGGGCCAGGC GGCCATCGAT AGCCCCAACC TCTGCCTGCG TCTCCGGTGC	180
60	TGCTACCGCA ATGGGGTCTG CTACCACCAAG CGTCCAGACG AAAACGTGCG GAGGAAGCAC	240

	ATGTGGCGC TGGTCTGGAC GTGCAGCGGC CTCCCTCC TGAGCTGCAG CATCTGCTTG	300
5	TTCTGGTGGG CCAAGCGCCG GGACGTGCTG CATATGCCCG GTTTCCTGGC GGGTCCGTGT	360
	GACATGTCCA AGTCCTGCTC GCTGCTCTCC AAGCACCGAG GGACCAAGAA GACGCCGTCC	420
	ACGGGCAGCG TGCCAGTCGC CCTGTCCAAA GAGTCCAGGG ATGTGGAGGG AGGCACCGAG	480
10	GGGGAAGGGG CGGAGGAGGG TGAGGAGACA GAGGGCGAGG AAGAGGAGGA TTAGGGAGT	540
	CCCCGGGGGA CTGGTCAATA CAGATACGGT GGACGGAAAA AAAAAAAAAA AAAAAAAA	598

15

## (2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
- |    |                            |
|----|----------------------------|
| 20 | (A) LENGTH: 454 base pairs |
|    | (B) TYPE: nucleic acid     |
|    | (C) STRANDEDNESS: double   |
|    | (D) TOPOLOGY: linear       |

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

	ATGGAGGCTG TTTTACAGT TTTTTTTTTT GTTGTGTTT TGTTTTTAAA GAATACAGAA	60
	GGAGCCAAGC TTTTTTGACAC TTTGTATCCA GCTGCAAGCT CAGGGCAGAG TCAAGGGCT	120
30	GGGTTGGAAA AACCTGACTC ACAGGAATGC ATAATTGACC CTTGCAGCTA CCCAATAGCC	180
	CTTGGAGCTG GCACGTGAAAC AGGCTGCAAG ATTGTACTGC CTTAAAAACA CAAGGCCCTC	240
35	TAGGCCTGGC AGGGATGTCC CTGTGCCAG CACTGGGGC TCGAAGACTG GTTTCTAGCA	300
	CTACCGGTCA CGGCCATGTC GTCCTAGAAG GGTCCAGAAG ATTATTITAC GTTGAGTCCA	360
	TTTTTAATGT TCTGATCACC TGACAGGGCA CCCCAAACCC CCAACTCCCA ATAAAAGCCG	420
40	TGACGTTCGG ACAAAAAAAAAA AAAAAAAAAA AAAA	454

45

## (2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
- |    |                            |
|----|----------------------------|
| 50 | (A) LENGTH: 425 base pairs |
|    | (B) TYPE: nucleic acid     |
|    | (C) STRANDEDNESS: double   |
|    | (D) TOPOLOGY: linear       |

## 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

	GCTAAAGGCC ATTCCCTCCG CAGGGCATTT GGCCTCGGGT GGGAGGGAA AACGCATCTT	60
	GTAAATTATT TTTAATCTTA TTTATTGTAC ATACCTGGGG CAGGGCTTG GGGAGGTGGA	120
60	GGGGGRAGAA GGGTCCCCTC TCTCTGCCCT TCCCACTCCT TTTCTACGGC GATTGTCTG	180

	TGTCTGGCCC CCACCCACTG MCCATCCCCC ATTGTTGTCT GGATGTGGTT CTATTTTTA	240
5	TCGGTCTCCT TTCCCCCTCCT CCCC GTTYTC GCCCCCGMCC CACCCCTGC TCCC ACTACC	300
	CTTTGTCTCT TGCTCTTTCTG TGGGYTTCTG TACA ACTCAA CTTG TATA CACA CTGTGTACAC	360
	ACA ACCAGYC WAACGAAAA CCCAACGGCA AACACTTAA AAAAAAAA AAAAAACTGG	420
10	GGGGT	425

## 15 (2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2471 base pairs
  - (B) TYPE: nucleic acid
  - 20 (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

25	GGCACGAGTA TGGCTTCCCG TGGACTCAGC CTCCTCCCCG ANT CCTGGCA CGAGGGGGCT	60
	TCGGTCTGTG GCTTCCGTG GCTGACGTCA TCTGGAGGAG ATTTGCTTTC TTTTCTCCA	120
30	AAAGGGGAGG AAATTGAAAC TGAGTGGCCC ACGATGGAA GAGGGGAAAG CCCAGGGTA	180
	CAGGAGGCCT CTGGGTGAAG GCAGAGGCTA ACATGGGTT CGGAGCGACC TTGGCCGTIG	240
	GCCTGACCAT CTTTGTGCTG TCTGTCGTCA CTATCATCAT CTGCTTCACC TGCTCCTGCT	300
35	GCTGCCCTTA CAAGACGTGC CGCCGACCAAC GTCCGGTGT CACCA CACCA ACATCCACCA	360
	CTGTGGTGCA TGCCCCCTAT CCTCAGCCTC CAAGTGTGCC GCCCAGCTAC CCTGGACCAA	420
40	GCTACCAGGG CTACCA CACCA ATGCCGCCTC AGCCAGGGAT CCCAGCAGCA CCCTACCCAA	480
	TGCAGTACCC ACCACCTTAC CCAGCCCAGC CCATGGGCC ACCGGCCTAC CACGAGACCC	540
	TGGCTGGAGA GCAGCCGGCG CCTACCCCGC CAGCCAGGCT CCTTACAACC CGGCCTACAT	600
45	GGATGCCCG AAGGGGGCCC TCTGAGCATT CCCTGGCCTC TCTGGCTGCC ACTTGGTTAT	660
	GTTGTGTGTG TCGGTGAGTG GTGTGAGGC GCGGTTCTT ACGCCCCATG TGTGCTGTGT	720
50	GTGTCCAGGC ACGGTTCCCTT ACGCCCCATG TGTGCTGTGT GTGTCTGCTC TGTATATGTG	780
	GCTTCCCTCTG ATGCTGACAA GGTGGGGAAC AATCCTTGCC AGAGTGGCT GGGACCAGAC	840
	TTTGTCTCTC TCCTCACCTG AAATTATGCT TCCTAAAATC TCAAGCCAAA CTCAAAGAAT	900
55	GGGGTGGTGG GGGGCACCC GTGAGGTGGC CCCTGAGAGG TGGGGGCCTC TCCAGGGCAC	960
	ATCTGGAGTT CTTCTCCAGC TTACCCTAGG GTGACCAAGT AGGGCCTGTC ACACCAGGGT	1020
60	GGGGCAGCTT TCTGTGTGAT GCAGATGTGT CCTGGTTTCG GCAGCGTACC AGCTGCTGCT	1080

	TGAGGCCATG GCTCCGTCCC CGGAGTTGGG GGTACCCGTT GCAGAGCCAG GGACATGATG	1140
	CAGGCGAAGT TGGGGATCTG GCCAAGTTGG ACTTTGATCC TTTGGCAGA TGTCCTATTG	1200
5	CTCCCTGGAG CCTGTCATGC CTGTTGGGA TCAGGCAGCC TCCTGATGCC AGAACACCTC	1260
	AGGCAGAGCC CTACTCAGCT GTACCTGTCT GCCTGGACTG TCCCCGTGCC CCCCATCTCC	1320
	CCTGGGACCA GCTGGAGGGC CACATGCACA CACAGCCTAG CTGCCCCAG GGAGCTCTGC	1380
10	TGCCCTTGCT GGCCCTGCC TTCCCCACAGG TGAGCACGGC TCCCTGTCCAC CAGCACACTC	1440
	AGTTCTCTTC CCTGCAGTGT TTTCATTTA TTTTAGCCAA ACATTTGCC TGTTTCTGT	1500
15	TTCAAACATG ATAGTTGATA TGAGACIGAA ACCCCTGGGT TGTGGAGGG AATTGGCTCA	1560
	GAGATGGACA ACCTGGCAAC TGTGAGTCCC TGCTTCCCGA CACCAGCCTC ATGGAATATG	1620
	CAACAACCTCC TGTACCCAG TCCACGGTGT TCTGGCAGCA GGGACACCTG GGCAATGGG	1680
20	CCATCTGGAC CAAAGGTGGG GTGTGGGCC CTGGATGGCA GCTCTGGCCC AGACATGAAT	1740
	ACCTCGTGT CCTCCCTCCCT CTATTACTGT TTCACCAAGAG CTGCTTCTGC TCAAATCTGT	1800
25	TGTGTTCTG AGTCTAGGGT CTGTACACTT GTTTATAATA AATGCAATCG TTTGGAAAAA	1860
	AAAAAAAAAA AAACCTCGTAG GGGGGGCCCG TACCCAATGG GCYCMMARAT AGTAGARWAC	1920
30	RAAAAYAMCA ANTGCAACCA AAGAGGGGCC AGGGGANITT TAAGAGGGCC CCCTTTGGG	1980
	GGNATCCANT TTAGCCGGGG TTNTTAAGGG AAGTTGCNTG GCGGGGGTTA GGGCCCSGTT	2040
	KYTWCCTCCA ACCAAGGGTT YTGYTGGTTA GGCCGGGTTG GGCCCMATGG GCTGGGCTGG	2100
35	GTAAAAGTGGT GGGTMAYTGC MATTGGTAG GGTGCTGCTG CCATTCTGG CTGAGGCGGC	2160
	ATGGTGTGGT AGCCCTGGTA GCTTGGTCCA GGGTAGCTGG GCGGCACACT TGGAGGCTGA	2220
	GGATAAGGGG CATGCACCCA CAGTGGTGA TGTGGTGGT GTGACAACCG GACGTGGTCG	2280
40	GCGGCACGTC TTGTAAAGGC AGCAGCAGGA GCAGGTGAAG CAGATGATGA TAGTGACGAC	2340
	AGACAGCACA AAGATGGTCC AGCCAACGGC CAAGGTCGCT CCGAACCCCA TGTAGCCTC	2400
45	TGCCTTCACC CAGAGGCCCTC CTGTACCCCT GGGCTTCCCC CTCTTCCCAT CGTGGGCCAC	2460
	TCACTCGTGC C	2471

50

(2) INFORMATION FOR SEQ ID NO: 42:

55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2659 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

	GGCACGAGCT TTTCTCTAGA GTCTGAAAGA TGCTAGAAAAG AAATAAAAATT TAACCTTACTT	60
5	AAGAGAATTA TGGATCTTTT ATTAATAAAA ATTAACCTGA TGATTTGAAC TAACAGITAT	120
	GATAATTCTG GTATTTATAG CTTTTTTAT TCCCCTGCAG AAAACCATAG GCAAAATTGC	180
10	AACATGCTTG GAATTGCGAA GTGCAGCTTT ACAGTCACA CAGTCTCAAG AAGAATTAA	240
	ACTGGAGGAC CTGAAGAACG TAGAACCAAT CCTAAAGAAT ATTCTTACAT ATAATAAAGA	300
	ATTCCCATTG GATGTTCAGC CTGTCCCATT AAGAAGAATT TTGGCACCTG GTGAAGAAGA	360
15	GAATTTGGAA TTTGAAGAAG ATGAAGAAGA GGGTGGTGCT GGAGCAGGTC TCCTGATTCT	420
	TTCCCTGCTAG AGTICCCGGT ACTTTATTAC CAAGGTTGCC ATCGGAACCA GGAATGACAT	480
	TACTCACTAT CAGAATTGAG AAAATTGGTT TGAAAGATGC TGGGCAGTGC ATCGATCCCT	540
20	ATATTACAGT TAGTGTAAAG GATCTGAATG GCATAGACTT AACTCCTGTG CAAGATACTC	600
	CTGTGGCTTC AAGAAAAGAA GATACATATG TTCATTAA TGIGGACATT GAGCTCCAGA	660
25	AGCATGTTGA AAAATTAAACC AAACGTGCAG CTATCTCTT TGAATTCAA CACTACAAGC	720
	CTAAAAAAAG GTTTACCAAGC ACCAAGTGTGTT TTGCTTTCAT GGAGATGGAT GAAATTAAAC	780
	CTGGGCCAAT TGTAATAGAA CTATACAAGA AACCCACTGA CTTAAAAGA AAGAAATTGC	840
30	AATTATTGAC CAAGAAACCA CTTTATCTTC ATCTACATCA AACTTTGCAC AAGGAATGAT	900
	CCTGACATGA TGAACCTGGA ACPTCTGTGA ATTTTACCAAC TCAGTAGAAA CCATCATAGC	960
35	TCTGTGTAGC ATATTCACCC TTCAACAGGC AGGAAGCAAG CCGTACCCAG ACCAGTAGGC	1020
	CGGACGGAGT CAAATGCAA GCTGTACCAAC AGAATTCAAGA GTCCAGCACA TCACACTGAC	1080
	GTATAGGACT CCTTGGGATA CAGGTTTATT GTAGATTTG AAACATGTTT TTACTTTCT	1140
40	ATTAATGTG CAATTAATAG TCTATTTCT AATTTACCAAC TACTCCTACC CTGCTTCCTG	1200
	GAACAATACT GTTGTGGGTA GGATGTGCTC ATCTTCAGAC TTAATACAGC AATAAGAATG	1260
45	TGCTAGAGTT TACACATCTG TTCACTTTG CTCCAATATG CTCTTTGAC TTAACGTCAA	1320
	GCTTTGGGTT GATGTGGGTA GGGTAGTGTGTC AAACGTGTTT GAGAGGAATG GGACCAAGTTC	1380
	TGCTGCCTAA GAAGGTCTGT CTGGATGTTT ATAGGCAGCA CCTCTGAAGT GGCCTAAATT	1440
50	CACCTGATC TGATAGTTT CCTGCTTAGA AAGTGTGCCT TGGCCAGATC AGTATCCCAC	1500
	ATGGGAGTGT TCCCTAGGTT GTAGCTGTGA TTGTTCCAG ATGACCAGAT TGTTTTCTG	1560
55	AAAATGAGCA TATTTTAGT CATGTCGATT AGCTGTCTT CTACATCACA TTGTTACTCT	1620
	TTCTGTGATGAT GATTCTAGGG TTAACATTGG AACCATCTCA AAATAATTAC AAAGTTTTAG	1680
	ATGGGTTTAC AATGTCTTCT AAACAATGTA ATCTAAAAAT AATTGAGTCA GATGCTAACG	1740
60	AGATACTGCA GGCATAACTG CTGTTTTCT GACAACGTAT TGTGAAACCT TAAAACCTGC	1800

	ATACCTCTTC TTACAGTGAG GAGTATGCAA AATCTGGAAA GATATTCTAT TTTTTTTATA	1860
5	TAGGTAGATA GGATGCCAT TTATTCCTA TTTAGATATA CTGACATTCA TCCATATGAA	1920
	AATATGCAGG TCATTAGCTT ACTATAATT ACTTTGACT TAATGGGCA TAAATAAAC	1980
	TTTCATAGTA CACATGAGGT GGATATTGA TACACAGAAC ATTTGCGGT GGCTTCTGT	2040
10	GGGTTAGATG TAAAGCCCAC ATATTTAAT ATTCACTATT TTAAATGAGC AATGCATGAG	2100
	GGGAATGCAG TGTCAGTACC TGGCCTATT TTAAACTAGT GTAATCACCC TAGTCATACC	2160
15	ATTCACTATG TTTGCTTTT AAAATAAGTA ACCACAATTA AGTTGTTGTA GCCCTTGCAC	2220
	TTCAAGAGAT CTAGTCCTTA CTTTCAGTTG TCTGTTAGGT CCATTCTGTT TACTAGACGG	2280
	ATGTTAATAA AAACATATGCG AGCCTGGAAT GGAATTCTCC AGCCAAATT TAGTCTTGTC	2340
20	CTCTCCATCT TGATTGGATT AATTCCAAT TCTAAATGA TTCAGTCCAC AATAGCTCTA	2400
	GGGGATGAAG AATTGCCCT ACTTTGCCA GTTCTTAAGA CTGTGAGTTG TCAAATCCCT	2460
25	AGACTGTAAG CTCTCAAGG AGCAAGAGGC GCATTTCTC CGTGTCTGTT AATTCTCTA	2520
	AGGTGTTGG CAGCACTCTG TACCTGTGG AGTACTCAGT ACCTTTGTT TGATGTTGCT	2580
	GACAAGACCT GAAAAAAAT CCCTAAAAA AAAAACCAT TAAAGTGTAG CAAAACCGAA	2640
30	AWAAAAAAA AAAAAAAA	2659

## 35 (2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1635 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

45	CGAGGAGGTC ATGAAACAAGG AGGGGGAGA GGTGGACGTG GTGGCTATGA CCATGGTGGC	60
	CGAGGGGGAG GAAGAGGAAA TAAGCATCAA GGAGGCTGGA CAGATGGAGG GAGTGGTGGA	120
	GGAGGTGGCT ACCAAGATGG TGGTTATCGA GATTCAAGTT TCCAGCCAGG TGGCTATCAT	180
50	GGTGGCCACA GCAGTGGTGG CTATCAAGGC GGAGGTTATG GTGGCTTCCA AACATCTCT	240
	TCATATACAG GAAGTGGATA CCAGGGTGGT GGCTACCAGC AGGACAATAG ATACCAAGAT	300
55	GGCGGGCACC ATGGTGATCG TGGTGGTGGT CGTGGTGGGC GAGGTGGTCG TGGAGGCCGA	360
	GGTGGTGTG CAGGCCAGGG AGGAGGCTGG GGAGGAAGAG GGAGCCAGAA TTATCACCAA	420
60	GGGGTCAAT TTGAACAGCA TTTCCAGCAT GGAGGTTATC AGTATAATCA TTCTGGATTT	480

	GGACAGGGAA GACATTACAC TAGTTGAGGC TACCGAACCT TACATTTTGC TAGAGCTCAA	540
	GTAATAGAAA CTTAGTTCA GAATCCTGAA TTCAGCACCT ATTTGAATT AATGTGAGAC	600
5	CACAGGTGGC AGGCAGATTG CTGCTTGGCA TAAGCATTG TAGGTCTTCA TTCAATTCTG	660
	TTAGATTTTT TTATTGGACT TACATAATGC CGTTTATTG AGAAACACAT AACATCTCTC	720
10	CTTTCTATGA AAAATTTTTT AAAAGGTGGT TAAAATTGCC TTTAATTGCC CAGTAGACTA	780
	ATTCACAGT CAGAACATGC AAACTTTTT GAAGAAATTA CTTGAATAAG TAGTTTCAT	840
	GTTCATCAA TGCAGTTTG AAAATGAGGA TTCACCTAGA CTTTTTAGA TTTACTACYA	900
15	GGAAACCTTC CYCATATGAA TAACCATTAA TATGTGTTT GCTTAAAGTA TTCCAATGCC	960
	TATTTTCCAA GCACAGTTCT GCCCCCCGGT TGACTTTAT GCCACGTGTG CTTCATGATG	1020
20	GAACTTTTAG GTCAGTTCT ATTAAATGAG CTCTTYTGCA GATAGCACAT TCAGTAGCCT	1080
	TATTTTGTIG ATGGAATACT GTATCATATG CTCAACTCTG AAAACCTTGA ACACGGCCAA	1140
	AATCCATAAA GATTATAAAA GCAAACATAAG TTGTGAAGCT ATAGTACATG TAGGCATTTA	1200
25	GTAAAGTATA GCAATTCAA CTGACCTGCA TCCATCCAAA ACAAAATTCT CCTTCACCT	1260
	TATTTTACT TGAAATTGCA TAGAAGAAAT AGCAAACCGA AATTTGTTT ATGCATGAGT	1320
30	TAATACCACT GGCTCAGCAA ATACAAGTTA GTTTGCTTTA AGCAGGTAAC TTTTTTGTAA	1380
	ATGGAAGAAA TGCACACAA AGTTAAGACA GATTTTGCT AAGTGCAGGA GGCCCTTTAT	1440
	TATTGCTGCA GAAAACAAAA GCCTGGCTGA GTTGATGTTT TACATTCTCC CTTACTGAAA	1500
35	TCTACATGAC ATGATGCTTC TTGCTGGGT TTTGTACATG TAAACATTGT CAAGCTGTGA	1560
	AAGAAAATGG CTGGAGGTGT GCTTGTGTG AAAGGTGAGC ACTGAAAGTA TCTGTTAAGT	1620
	TCTCCNGAAA AAAAA	1635
40		

## (2) INFORMATION FOR SEQ ID NO: 44:

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 780 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

55	AACATGGTCA TGTCTTTAG TTTCATTATT TTCTACTCC TTGTATGTCA AGAAATTACA	60
	TTTTGCATGT CTTATGGAGA TGCTGTTAAT TGCTTCAGTG AGTGTGTTTC TAATCTGCAG	120
	ACCATTACAA TTTCCTGTT GCAGCATGCT GTGTGCAAAC AYTCAGTAAT TTGGAGTATT	180
60	CAATTATTTG TTAGGGCTCT TCCTATTCC AAATGTGCTG AATTGTCTAT TGATGGGATT	240

	TTCAGATCTT TTCATGAGAA CTGGAAATGT AGCTGGTGG CACCTACCTA GGTTGCTACG	300
5	TAGTGAGTAG ACTTTCTCTT GGGTATAGTA AGCCTCAGAC AGCTTTCACT TTTATCTACT	360
	TTACTTGTGG AAATAAAAACA GTCAATTGT TCTGAAAGAA TAAGATAGCT TTCTGTAGAG	420
	AAGGAATTCC TACCTCTAAA AGCTGCCTTG AGAACTCAGA ACTGGCAGTT TTCTGAGGTG	480
10	ATTTTTAAAT TTCAGTATTA GGGAGAGTCC AGCATTGCT GACACAGATT CTACATAACT	540
	AATGTATGAT AGCAAATGCA AAACTATTAT AATGTGGTGT ATCTTGCGCA TACACAGGT	600
	AGAACAAAGTA GACTCTGGCA GCAGATCTCC AGAGACCAA GTTTAGGTC TCATAGTGT	660
15	TTTGAAGTAG TTATRACTCCT GGCTTAAGTA GTTTAGTGC C TGGGAGAAC CATTACTGAA	720
	AAGCATTAA CTTAAAAAAA AAAAAAAAAA AAAACTGAAA AGTAGTGAA TACAGAATAG	780
20		

## (2) INFORMATION FOR SEQ ID NO: 45:

## 25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

	GCGAAGCAGC TGAAGCCGCC GCCGCGCAGA ATCCACGCTG GCTCCGTGCG CCATGGTCAC	60
35	CCACAGCAAG TTTCCCGCCG CCGGGATGAG CCGCCCCCTG GACACCAGCC TGCGCCTCAA	120
	GACCTTCAGC TCCAAGAGCG AGTACCAAGCT GGTGGTGAAC GCAGTGCAGCA AGTGCAGGAG	180
	AGCGGTTCT ACTGGAGCCG AGTGACCGGC GGCGAGGCGA ACCTGCTGCT CAGTGCCGAG	240
40	CCCGCCGGCA CCTTTCTGAT CCGCGACAGC TCGGGACCAG CGCCACTTCT TCACGCTCAG	300
	CGTCAAGACC CAGTCTGGGA CCAAGAACCT GCGCATCCAG TGTGAGGGGG GCAGCTTCTC	360
45	TCTGCAGAGC GATCCCCGGA GCACGCAGCC CGTGSCCCGC TTCGACTGCG TGCTCAAGCT	420
	GGTGCACCAC TACATGCCGC CCCCTGGAGC CCCCTCCTTC CCCTCGCCAC CTACTGAACC	480
	CTCCCTCCGAG GTGCCCGAGC AGCCGCTG CCAAGCCACTC CCTGGGAGTC CCCCCAGAAG	540
50	AGCCTATTAC ATCTACTCCG GGGGGAGAA GATCCCCCTG GTGTTGAGCC GGCCCTCTC	600
	CTCCAACGTG GCCACTCTTC AGCATCTCTG TCGGAAGACC GTCAACGGCC ACCTGGACTC	660
55	CTATGAGAAA GTCACCCAGC TGCCGGGGCC CATTGGGAG TTCCTGGACC AGTACGATGC	720
	CCCGCTTTAA GGGGTAAAGG CGCAGAAAGGG CATGGGTGG GAGAGGGGAC GCAGGCCCCCT	780
60	CTCCCTCCGTG GCACATGGCA CAAGCACAAG AAGCCAACCA GGAGAGAGTC CTGTAGCTCT	840

	GGGGGGAAAG AGGGCGGACA GGCCCCCTCCC TCTGCCCTCT CCCTGCAGAA TGTGGCAGGC	900
	GGACCTGGAA TGTGTTGGAG GGAAGGGGAA GTACCACCTG AGTCTCCAGC TTCTCCGGAG	960
5	GASCCAGCTG TCCTGGTGGG ACGATAGCAA CCACAAGTGG ATTCTCCTTC AATTCCCTCAG	1020
	CTTCCCCCTCT GCCTCCAAAC AGGGGACACT TCGGGAATGC TGAACTAATG AGAACTGCCA	1080
10	GGGAATCTTC AAACCTTCCA ACGGAACCTTG TTTGCTCTTT GATTGGTTT AAACCTGAGC	1140
	TGGTTGTGGA GCCTGGAAA GGTGGAAGAG AGAGAGGTCC TGAGGGCCCC AGGGCTGCCG	1200
	GCTGGCGAAG GAAATGGTCA CACCCCCCGC CCACCCCAAG CGAGGATCCT GGTGACATGC	1260
15	TCCTCTCCCT GGCTCCGGGG AGAAGGGCTT GGGGTGACCT GAAAGGGAAC CATCCTGGTG	1320
	CCCCACATCC TCTCCTCCGG GACAGTCACC GAAAACACAG GTTCCAAAGT CTACCTGGTG	1380
20	CCTGAGAGCC CAGGGCCCTT CCTCCGTTTT AAGGGGAAAG CAACATTTGG CACGAGATGG	1440
	GCTGGTCAGC TGGTCTCCTT TTCCTACTCA TACTATACTT TCCTGTACCT GGGTGGATGG	1500
	ACCGGGAGGA TGGAGAGACG GGACATCTT CACCTCAGGC TCCTGGTAGA GAATACAGGG	1560
25	GATTCTACTC TGTGCCCTCCT GACTATGTCT GGCTAAAGAGA TTCCGCTTAA ATGCTCCCTG	1620
	TCCCCTGGAG AGGGACCCAG CATAGGAAAG CCACATACTC AGCCTGGATG GGTGGAGAGG	1680
30	CTGAGGGACT CACTGGAGGG CACCAAGCCA GCCCACAGCC AGGGAAAGTGG GGAGGGGGC	1740
	GGAAACCCAT GCCTCCCAGC TGAGCACTGG GAATGTCAGC CCAGTAAGTA TTGGCCAGTC	1800
	AGGCGCCTCG TGGTCAGAGC AGGCCACCA GGTCCCCTG CCCCGAGCCC TGCACAGCCC	1860
35	TCCCTCTGC CTGGGTGGGG GAGGCTGGAG GTCATGGAG AGCCTGGACT GCTGCCACCC	1920
	CGGGTGCTCC CGCTCTGCCA TAGCACTGAT CAGTGACAAT TTACAGGAAT GTAGCAGCGA	1980
40	TGGAATTACC TGGAACAGTT TTTTTTTTTT GTTTTTGTTT TTGTTTTTGT GGGGGGGGGC	2040
	AACTAAACAA ACACAAAGTA TTCTGTGTCA GGTATGGGC TGGACAGGGC AGTTGTGTGT	2100
	TGGGGTGGTT TTTTTCTCTA TTTTTTTGTT TGTTTCTTGT TTTTTAATAAA TGTTTACAAT	2160
45	CTGCCCTCAAT CACTCTGTCT TTTATAAAGA TTCCACTCCA GTCCTCTCTC CTCCCCCTA	2220
	CTCAGGGCCCT TGAGGCTATT AGGAGATGCT TGAAGAACTC AACAAAATCC CAATCCAAGT	2280
	CAAACTTTGC ACATATTTAT ATTTATATTC AGAAAAGAAA CATTTCAGTA ATTTATAATA	2340
50	AAGAGCACTA TTTTTTAATG AAAAAAAA AAAAAAAA	2378

55

(2) INFORMATION FOR SEQ ID NO: 46:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1772 base pairs
- (B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

5	TCGACCCACG CGTCCCCAG GATCCCCAGC CGGGTCCCAA GCCTGTGCCT GAGCCTGAGC	60
10	CTGAGCCTGA GCCGAGCCGG GAGCCGGTCG CGGGGGCTCC GGGCTGTGGG ACCGCTGGC	120
15	CCCCAGCGAT GGCGACCTG TGGGGAGGCC TTCTTCGGCT TGGCTCCTTG CTCAGCCTGT	180
20	CGTGCCTGGC GCTTTCGGTG CTGCTGCTGG CGCACTGTCA GACGCCGCCA AGAATTTCGA	240
25	GGATGTAGA TGTAAATGTA TCTGCCCTCC CTATAAAAGAA AAATTCTGGG CATAATTATA	300
30	ATAAGAACAT ATCTCAGAAA GATTGTGATT GCCTTCATGT TGTGGAGCCC ATGCCTGTGC	360
35	GGGGGCCTGA TGTAGAACCA TACTGTCTAC GCTGTGAATG CAAATATGAA GAAAGAACCT	420
40	CTGTCACAAT CAAGGTTACC ATTATAATTIT ATCTCTCCAT TTTGGGCCCTT CTACTTCCTG	480
45	ACATGGTATA TCCTACTCTG GTTGAGCCCA TACTGAAGAG GCGCCTCTTT GGACATGCAC	540
50	AGTTGATACA GAGTGTGAT GATATTGGGG ATCACCAAGCC TTTTGCAAAT GCACACGATG	600
55	TGCTAGCCCG CTCCCGCAGT CGAGCCAACG TGCTGAACAA GGTAGAATAT GGCACAGCAG	660
60	CGCTGGAAGC TTCAAGTCCA AGAGCAGCGA AAAGTCTGTC TTTGACCCGC ATGTTGTCCT	720
65	CAGCTAATTG GGGATTGAA TTCAAGGTGA CTAGAAAGAA ACAGGCAGAC AACTGGAAAG	780
70	GAACGTACTG GTTTTTGCTG GGTTCATTT TAATACCTTG TTGATTICAC CAACTGTTGC	840
75	TGGAAGATTC AAAACTGGAA GKAAAAACTT GCTTGATTTT TTTTTCTTGT TAACGTAATA	900
80	ATAGAGACAT TTTTAAAGC ACACAGCTCA AAGTCAGCCA ATAAGTCTTT TCCTATTGT	960
85	GACTTTTACT AATAAAAATA AATCTGCCTG TAAAATAAT TAAAAAATCC TTTACCTGGA	1020
90	ACAAGCACTC TCCTTTTCAC CACATAGTTT TAACTTGACT TTCCAAGATA ATTTTCAGGG	1080
95	TTTTTGTGT TGTTGTTTT TGTTGTTTG TTTTGGTGGG AGAGGGGAGG GATGCCCTGG	1140
100	AAGTGGTTAA CAACTTTTTT CAAGTCACCT TACTAAACAA ACTTTTGAA ATAGACCTTA	1200
105	CCTTCTATTT TCGAGTTCA TTTATATTTT GCAGTGTAGC CAGCCTCATC AAAGAGCTGA	1260
110	CTTACTCATT TGACTTTGC ACTGACTGTA TTATCTGGGT ATCTGCTGTG TCTGCACTTC	1320
115	ATGGTAAACG GGATCTAAAA TGCCTGGTGG CTTTCACAA AAAGCAGATT TTCTTCATGT	1380
120	ACTGTGATGT CTGATGCAAT GCATCCTAGA ACAAACTGGC CATTTGCTAG TTTACTCTAA	1440
125	AGACTAAACA TAGTCTTGGT GTGTGTGGTC TTACTCATCT TCTAGTACCT TTAAGGACAA	1500
130	ATCCTAAGGA CTGGACACT TGCAATAAG AAATTTTATT TTAAACCCAA GCCTCCCTGG	1560
135	ATTGATAATA TATACACATT TGTCAGCATT TCCGGTCTGTG GTGAGAGGCA GCTGTTTGAG	1620
140	CTCCAATGTG TGCAGCTTG AACTAGGGCT GGGTTGTGG GTGCCCTTC TGAAAGGTCT	1680

AACCATTATT GGATAACTGG CTTTTTTTCT TCCTCTTGG AATGTAACAA TAAAAATAAT	1740
TTTTGAAACA TCAAAAAAAA AAAAAAAA AA	1772

5

## (2) INFORMATION FOR SEQ ID NO: 47:

10

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1107 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

CGGGCGAGAA GGGCAGACGG GACATGCAGC CTCTTCGCC TGAGCCCCGG AAGTGATGTG	60
GCTGCGGCAT CGCGGCCTCG CTATGTCTGC CATTTCAT TTTCAGAGTC TATTGACTGT	120
AATCTTGCTG CTTATATGTA CCTGTGCTTA TATTCCATCC TTGGCACCCA GCCTCCTGG	180
CAGAAATAAA ACTGGATTGT TGGGTATATT TTGGAAGTGT GCCAGAATTG GTGAACGGAA	240
GAGTCCTTAT GTTGCAGTAT GCTGTATAGT AATGCCCTTC AGCATCCCT TCATACAGTA	300
GCTGGGAAA ATGCCAGAAT GTAGTTGCCA TCAGATTGTA TTGTGAACAA GGACTGACTG	360
CAGAAAATAA TGGAAAGGAT GTTTAACCTCT TTTATCTCCG AACATTGAAT GAGATAAATT	420
TCCAGATGCT GTTCTCTATT TTAATGTTAT TGGACCAATG TTCTGTATAA ACAATTAAAGA	480
TGTAACCATT TAATAGTCTG TAACAATCAA CCTCAGTACT GTCACTACAA TATTACATT	540
TGCAAATGTT ATTCTGTTGT ATCAGATACA AAATTTAGT GAGGTATCTC TAAGGCACAT	600
AGTAGAAAAC AAAATTGGTT AATTACTCAA GTTCCCTTCA CTGTGATTG GAAATGATT	660
AATCTTTATA GAATGAGAAC CTTTTTGGA CTAGCTTTT TATTAAATG GCTCAATTG	720
TGTTGATAAG GATTGCATTA ATATTTATA GTGCTTGCTT TTCTCTGGG CACACCATT	780
TGATCATTAA CCAGAGTACC TCTACTCTTA GCAAACCTCA GTTATGACA AGTATTTAAA	840
ATATTTAAA CAAGCTTATG CAGTTCTTAA GGACGAAGGT AAATGAGATG TAACTTAAAA	900
ATAGTATTGG GAAAATGTTG ATAGTTAACAA TTAGTGGATT TAGACTAGCC AAATGACATA	960
GTAGGGCTCG AAACATCTTG TCAAGTATAT GTATTTGTG CATGAATTG TGCTGGAAAG	1020
CTGTCTTTCT CTGAAAAACA CAACGTTCTT AGAATGAAA GAACAATTAT AAAATAAAA	1080
55 AAAATTTAA AAAAAACTGG CGGGGGG	1107

60 (2) INFORMATION FOR SEQ ID NO: 48:

## (i) SEQUENCE CHARACTERISTICS:

- 5           (A) LENGTH: 805 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: double  
              (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

10	TGCAGAAGAG ATGGAGTTGC TGTIGGAAAA CTACTACCGA TTGGCTGACG ATCTCTCAA	60
	TGCAGCTCGT GAGCCTAGGG TGCTGATTGA TGATTACAAA AGTATTATTT TCATTAATCT	120
15	GGACAGCCAC CGAAACGTGA TGATGAGGTT GAATCTACAG CTGACCATGG GAACCTTCTC	180
	TCTTCGCTC TTTGGACTAA TGGGAGTTGC TTTTGGAAATG AATTGGAAAT CTTCCCTTGA	240
	AGAGGACCAT AGAATTTTTT GGCTGATTAC AGGAATTATG TTCATGGAA GTGGCCTCAT	300
20	CTGGAGGCGC CTGCTTTCAT TCCTTGGACG ACAGCTAGAA GCTCCATTGC CTCCTATGGT	360
	ATGAAGGATA TGGTTCACGG CGGTATTGTG GAAGGGTTAT GATCATGGC CCTAAAGTCA	420
25	GAGCGCCTGG GATTAAGTTG TCACAGGCAC TATGGCCCTT GCGAGTTGCT TTCTCAAAC	480
	TCCTTCAGTT TCCCTATCTG TCAGTTAAGT CGGTATTACC TGCTTCATAG GGTTATGGGA	540
	AGAATTAAAC AATATGTGTA AACCACTTAC TAGCACACTG CCTAACACAA TAAGTTAGAA	600
30	ATATAATTIG TGTAGAACTC TGACAACATA CAITAAACACA GATGTTAGTA ATTCTGGTAT	660
	AAGGTTTGTC ATAACCAAAT GGAAATGTAG GAAACATTAA TAATGTTCTT AAAAGATAGR	720
35	AAATTACACT CCATTTCTT TGTACTTGAA GATGGCACCA CTGGAATAAA TACTTAAGAC	780
	ACTGAAAAAA AAAAAAAAAA AACTC	805

40

## (2) INFORMATION FOR SEQ ID NO: 49:

## (i) SEQUENCE CHARACTERISTICS:

- 45           (A) LENGTH: 1408 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: double  
              (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

50	TCATTATTTA TTCATGTGGC TGAAAGAGTA TATTAATTAT GTTTAGATTT TTGGAAAAAG	60
	TCTGAACAAA AAAAGGACCT ATACAGTGCT CAAACTATAT TTTTAAAAAT ACTATTTAT	120
55	TTTTACTCAC ATATGAAAAA AATGGCTGTA CTATCATGTT TACATACATA CTAACATTGG	180
	AAACAGAATA ACGAATTGTA TTTAAATTAT ATGAAGAACAA CACAAACATT AAAACACTGA	240
60	TTGGTTACAG AAACCAGACT TTGAGGAAAA AACATTAGCT ATAATTTCATTA	300

	AGAGCAGCAC CCTCTGAGAA TAATCAAAC T GATTAGTAAT APTCATCTAT ACTGCAAAAT	360
	AATATGTACA AAGGAAAGTT AGTGATTGTA CTGATTAT TACTTTTACC AAGCCATT TT	420
5	ATGTTCTCA CTCATGCCA AGAAATAAAA CATAATCTGA AGAAAAATAT GTCCCTTATTA	480
	TTATTACAA TAAAAGTTG GCTTTATTCT GCAAGCCTGG GCATATTGTA CAATTGGCAG	540
10	CACTTAACGG CTCAAGTGGA TCAATGTACC AGTTGATTG TGATCCACTG AATAGAACATCT	600
	CTCATCCATA TCTGGTGACC AGACTAACTC CATGGGAGCT GTGATAGACT GAACCATT TC	660
	TGTGGTATCC CTAGATCTCA CTAAATAAGA AAGACCCCTAC ACCAGAAAAT ATAGCAACTG	720
15	ATCTATCTAT AAATTACATC TATATGCTAG CTCTTAGTA TAAGTTGGAA AAAGGGCCC	780
	TTTCTTGAGC ACATGGATAA AAGTATTATT GTAGTCTAAA GATTGCTGGA TTGATATTGT	840
20	GTTGTTATAA TGAAGATAAG GTACACACTG AAACCACTGT CAGATTAAGA AACCTCCACA	900
	ACTTGTCTCA GTTCTTCAAA CAATGGAGCA AGTTCTTTT CTAGGCTGAC AATTAGTCCT	960
	GTATTGGCAC TGCTGCTGGC TATGAAACTC ACCACCAAAG GTAAACGATT AAATTGAACC	1020
25	ACCTGGTAGG TGTTATAGTA ACAGATGATA CTTTTATTT TGAAAAGTCC AAGTTTGCTT	1080
	CCTTGGCTCG TTGCAAGGGC AAAAGTGGAT AAGAAACCAAG GTCGCAAAGC ATGCTCTGGA	1140
	GCATTGTCAT TTGCCACTTT AATAACAGGT ACTCCATCTC TATCTGACAC AACAAATGGCA	1200
30	TGGAGCCCTT CAACACTTGG TAACTTTTA TACAAGAACATC GCTTTAGGTC ATCCGCCATG	1260
	ATGAACCCCCC TTCTCTCGCA GGATCAATCT CCACGCCCTGG GGTTTCTGGG CTGCCTGGTT	1320
35	CTCTCCCTG TCACCTCAGG GACAGCTTTA AAGACAGGTT CCTCCTCAAG CCACCGTCAC	1380
	ATGATTCACTG ACCTCGTCTG CGCTCCAG	1408

40

## (2) INFORMATION FOR SEQ ID NO: 50:

	(i) SEQUENCE CHARACTERISTICS:	
45	(A) LENGTH: 1813 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:	
	CATGGTGGGG CACGAGATGG CCTCTRACTC TTCWAACACT TCACTGCCAT TCTCAAACAT	60
	GGGAAATCCA ATGAACACCA CACAGTTAGG GAAATCACTT TTTCAGTGGC AGGTGGAGCA	120
55	GGAAGAAAGC AAATTGGCAA ATATTTCCCA AGACCAGTTT CTTTCAAAGG ATGCAGATGG	180
	TGACACGTTT CTTCTATATTG CTGTTGCCCA AGGGAGAAGG GCACCTTCCT ATGTTCTTGC	240
60	AAGAAAGATG AATGCACTTC ACATGCTGGA TATTAAGAG CACAATGGAC AGAGTGCCTT	300

	TCAGGTGGCA GTGGCTGCCA ATCAGCATCT CATTGTGCAG GATCTGGTGA ACATCGGGC	360
5	ACAGGTGAAC ACCACAGACT GCTGGGAAG AACACCTCTG CATGTGTGTG CTGAGAAGGG	420
	CCACTCCCAG GTGCTTCAGG CGATTAGAA GGGAGCAGTG GGAAGTAATC AGTTGTGGA	480
	TCTTGAGGCA ACTAACTATG ATGGCCTGAC TCCCCTCAC TGTGCAGTCA TAGCCCACAA	540
10	TGCTGTGGTC CATGAACCTCC AGAGAAATCA ACAGCCTCAT TCACCTGAAG TTCAGGAGCT	600
	TTTACTGAAG AATAAGAGTC TGGTTGATAC CATTAAAGTGC CTAATTCAA TGGGAGCAGC	660
	GGTGGAAAGCG AAGGATCGCA AAAGTGGCCG CACAGCCCTG CATTGGCAG CTGAAGAAC	720
15	AAATCTGGAA CTCATTGCC TCTTTTGGA GCTGCCAGT TGCCCTGTCTT TTGTGAATGC	780
	AAAGGCTTAC AATGGCAACA CTGCCCTCCA TGTGCTGCC AGCTTGAGT ATCGGTTGAC	840
20	ACAATTAGAT GCTGTCCGCC TGTGATGAG GAAGGGAGCA GACCCAAGTA CTCGGAACCTT	900
	GGAGAACGAA CAGCCAGTGC ATTGGTTCC CGATGGCCCT GTGGGAGAAC AGATCCGACG	960
	TATCCTGAAG GGAAAGTCCA TTCAGCAGAG AGCTCCACCG TATTAGCTCC ATTAGCTTGG	1020
25	AGCCTGGCTA GCAACACTCA CTGTCAGTTA GGCAGCCTG ATGTATCTGT ACATAGACCA	1080
	TTTGCCTTAT ATTGGCAAAT GTAAGTTGTT TCTATGAAAC AAACATATTT AGTICACTAT	1140
30	TATATAGTGG GTTATATTAA AAGAAAAGAA RAAAAATATC TAATIWCTCT TGGCAGATTT	1200
	GCATATTTCATACCCAGGTA TCTGGATCTA GACATCTGAA TTTGATCTCA ATGGTAACAT	1260
	TGCCCTCAAT TAACAGTAGC TTTTGAGTAG GAAAGGACTT TGATTTGTGG CACAAAACAT	1320
35	TATTAATATA GCTATTGACA GTTCAAAAGC AGGTAAATTG TAAATGTTTC TTTAAGAAAA	1380
	AGCATGTGAA AGGAAAAGG TAAATACAGC ATTGAGGCTT CATTGGCCT TAGCCCTGG	1440
40	GAGTTACTGG CGTTGGACAG GCTTCAGTCA TTGACTAGA TGAAAGGTGT CCATGGTTAG	1500
	AATTGATCT TTGCAAACGT TATATAATTG TTATTTTGT CCTTAAAAAT ATTGTACATA	1560
	CTTGGTTGT TACATGGTCA TATTGAAAT GTATAAGTCC ATAAAATAGA AAAGAACAAAG	1620
45	TGAATTGTTG CTATTTAAA AAATTTACA ATTCTTACTA AGGAGTTTT ATTGTGTAAT	1680
	CACTAAGTCT TTGAGATAA AGCAGATGGG GAGTTACGGA GTTGTCCCTT TACTGGCTGA	1740
50	AAGATATATT CGAATTGTA AGATGCTTT YCTCATGCAT TGAAATTATA CATTATTTGT	1800
	AGGAAATTGC ATG	1813

55

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 2070 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

	CCACCGGTCC GGAAGAGCGC GGCACCTTCGG CTGGCCGCTG GCTCGCTGGC CGCTCCTGGA	60
10	GGCGGCCGCG GGAGCGCAGG GGGCGCGCGG CCCGGGGACT CGCATTCCCC GGTTCCCCCT	120
	CCACCCACG CGGCCTGGAC CATGGACGCC AGATGGTGGG CAGTGGTGGT GCTGGCTGCG	180
	TTCCTCTCCC TAGGGGCAGG TGGGGAGACT CCCGAAGCCC CTCCGGAGTC ATGGACCCAG	240
15	CTATGGTTCT TCCGATTGTG GGTGAATGCT GCTGGCTATG CCAGCTTTAT GGTACCAGGC	300
	TACCTCTGG TGCAGTACTT CAGGCGGAAG AACTACCTGG AGACCGGTAG GGGCCTCTGC	360
20	TTTCCCTGG TGAAAGCTTG TGTGTTGGC AATGAGCCA AGGCCTCTGA TGAGGTTCCC	420
	CTGGCGCCCC GAACAGAGGC GGCAGAGACC ACCCCGATGT GGCAGGCCCT GAAGCTGCTC	480
	TTCTGTGCCA CAGGGCTCCA GGTGCTTAT CTGACTTGGG GTGTGCTGCA GGAAAGAGTG	540
25	ATGACCCGCA GCTATGGGC CACAGCCACA TCACCGGGTG AGCGCTTTAC GGACTCGCAG	600
	TTCCTGGTGC TAATGAACCG AGTGTGGCA CTGATTGTGG CTGGCCTCTC CTGTGTTCTC	660
30	TGCAAGCAGC CCCGGCATGG GGCACCCATG TACCGGTACT CCTTTGCCA GCCTGTCAA	720
	TGTGCTTAGC AGCTGGTCCC AATACGAAGC TCTTAAGTTC GTCACTTCTC CCACCCAGGT	780
	GCTGGCCAAG GCCTCTAAGG TGATCCCTGT CATGCTGATG GGAAAGCTTG TGTCTCGCG	840
35	CAGTAACGAA CACTGGGAGT ACCTGACAGC CACCCCTCATC TCCATTGGGG TCAGCATGTT	900
	TCTGCTATCC AGGGGACCAAG AGCCCCGCAG CTCCCCAGCC ACCACACTCT CAGGCCTCAT	960
40	CTTACTGGCA GGTTATATTG CTTTTGAACA GCTTCACCTC AAACCTGGCAG GATGCCCTGT	1020
	TTGCCTATAA GATGTCATCG GTGCAGATGA TGTTTGGGG TCAATTCTT CTCCCTGCCTC	1080
	TTCACAGTGG GCTCACTGCT AGAAAACAGGG GGCCTACTG GAGGGAACCC GCTTCATGGG	1140
45	GCGACACAGT GAGTTTGCTG CCCATGCCCT GCTACTCTCC ATCTGCTCCG CATGTGGCCA	1200
	GCTCTTCATC TTTTACACCA TTGGGCAGTT TGGGGCTGCC GTCTTCACCA TCATCATGAC	1260
50	CCTCCGCCAG GCCTTGCCA TCCCTCTTTC CTGCCTTCTC TATGCCACA CTGTCACTGT	1320
	GGTGGGAGGG CTGGGGGTGG CTGTGGTCTT TGCTGCCCTC CTGCTCAGAG TCTACGGCG	1380
	GGGCCGTCTA AAGCAACGGG GAAAGAAGGC TGTGCCCTGT GAGTCTCCTG TGCAGAAGGT	1440
55	TTGAGGGTGG AAAGGGCCTG AGGGGTGAAG TGAAATAGGA CCCTCCCACC ATCCCTTCT	1500
	GCTGTAAACCT CTGAGGGAGC TGGCTGAAAG GGCAAAATGC AGGTGTTTTC TCAGTATCAC	1560
60	AGACCAGCTC TGCAGCAGGG GATTGGGGAG CCCAGGAGGC AGCCCTCCCT TTTGCCCTAA	1620

	GTCACCCATC TTCCAGTAAG CAGTTTATTIC TGAGCCCCGG GGGTAGACAG TCCTCAGTGA	1680
	GGGGTTTGGG GGAGTTGGG GTCAAGAGAG CATAGTAGG TTCCACAGTT ACTCTTCCA	1740
5	CAAGTTCCCT TAAGTCTTGC CCTAGCTGTG CTCTGCCACC TTCCAGACTC ACTCCCCTCT	1800
	GCAAATACCT GCATTTCTTA CCCGGTGAG AAAAGCACAA GCGGTGTAAGG CTCCAATGCT	1860
	GCTTTCCAG GAGGGTGAAG ATGGTGTGT GCTGAGGAAA GGGGATGCAG AGCCCTGCC	1920
10	AGCACCACCA CCTCCATATGC TCCGGATCC CTAGGCTCTG TTCCATGAGC CTGTTGCAGG	1980
	TTTTGGTACT TTAGAAATGT AACTTTTGC TCTTATAATT TTATTTTATT AAATAAATT	2040
15	ACTGCAAAAA AAAAAAAAAA AAAAAAAAAA	2070

## 20 (2) INFORMATION FOR SEQ ID NO: 52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1426 base pairs
- (B) TYPE: nucleic acid
- 25 (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

30	CCCTCACTAA AGGAAACAAA AGCTGGAGCT CCACCGCGGT GGCGGCCGCT CTAGAACTAG	60
	TGGATCCCCC GGGCTGCAGG AATTGGCAC ACGGATCGGC GTCCGCAGCG GGCGGCTGCT	120
	GAGCTGCCTT GAGGTGCAGT GTTGGGGATC CAGAGCCATG TCGGACCTGC TACTACTGGG	180
35	CCTGATTGGG GGCTGACTC TCTTACTGCT GCTGACGCTG CTGGCCTTIG CCGGGTACTC	240
	AGGGCTACTG GCTGGGTGG AAGTGAGTGC TGGGTACCCC CCCATCCGCA ACGTCACTGT	300
40	GGCCTACAAG TTCCACATGG GGCTCTATGG TGAGACTGGG CGGCTTTCA CTGAGAGCTG	360
	CAGCATCTCT CCCAAGCTCC GCTCCATCGC TGTCTACTAT GACAACCCCC ACATGGTGCC	420
	CCCTGATAAG TCCCGATGTG CGTGGGCAG CATCCTGAGT GAAGGTGAGG AATGCCCTC	480
45	CCCTGAGCTC ATCGACCTCT ACCAGAAATT TGGCTTCAAG GTGTTCTCCT TCCCGAAC	540
	CAGCCATGTG GTGACAGCCA CCTTCCCCCT AACACCACCA TTCTGTCCCA TCTGGCTGGG	600
50	CTACCCGCCG TGTCCATCCT GCCTTGGACA CCTACATCAA GGAGCGGAAG CTGTGTGCCT	660
	ATCCTCGGCT GGSGATCTAC CAGGAAGACC AGAATCCATT TCATGTGCC ACTGGCACGG	720
	CCAGGGAGAC TTCTATGTGC CTGAGATGAA GGAGACAGAG TGGAAATGGC GGGGGCTTGT	780
55	GGAGGCCATT GACACCCAGG TGGATGGCAC AGGAGCTGAC ACAATGAGTG ACACGAGITC	840
	TGTAAGCTTG GAAGTGAGCC CTGGCAGCCG GGAGACTTCA GCTGCCACAC TGTACACCTGG	900
60	GGCGAGCAGC CGTGGCTGGG ATGACGGTGA CACCCGCAGC GAGCACAGCT AACAGCGAGT	960

	CAGGTGCCAG CGGCTCCCTCT TTTGAGGAGC TGGACTTGG AGGGCGAGGG GCCCTTAAGG	1020
5	GGAGTCACGG CTGGACCCCTG GGACTTGAGC CCCTGGGGGA CTACCAAGTG GCTCTGGGAG	1080
	CCCACTGCC C TGAGAAGGG CAAGGACTAA CCCATGCCCT GCACCCCTCC GCAGTGCAGT	1140
	TGCTGAGGAA CTGAGCAGAC TCTCCAGCAG ACTCTCCAGC CCTCTTCCTC CTTCCCTCTGG	1200
10	GGGAHGAGGG GTTCTGAGG GACCTGACTT CCCCTGCTCC AGGCCTCTTG CTAAGCCTTC	1260
	TCCTCACITGC CCTTTAGGCT CCCAGGGCCA GAGGAGCCAG GGACTATTTT CTGCACCAGC	1320
	CCCCAGGGCT GCGGCCCTG TTGTGTCTTT TTTTCAGACT CACAGTGGAG CTTCCAGGAC	1380
15	CCAGAATAAA GCCAATGATT TACTTGTAA AAAAAAAA AAAAAA	1426

20

## (2) INFORMATION FOR SEQ ID NO: 53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1720 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

30	GGCACGAGTG CGGGCCCAGC CTCTCCTCAC GCTCGGCCAG TCTCCGCCGC AGTCTCAGCT	60
	GCAGCTGCAG GACTGAGCCG TGCAACCGGA GGAGACCCCC GGAGGAGGCG ACAAACTTCG	120
35	CAGTGCAGCG ACCCAACCCC AGCCCTGGGT AGCCTGCAGC ATGGCCCAAGC TGTTCCCTGCC	180
	CCTGCTGGCA GCCCTGGTCC TGGCCCAGGC TCCTGCAGCT TTAGCAGATG TTCTGGAAGG	240
40	AGACAGCTCA GAGGACCGCG CTTTTCGGGT GCGCATCGCG GCGGACGCGC CACTGCAGGG	300
	CGTGCTCGGC GGCGCCCTCA CCATCCCTTG CCACGTCCAC TACCTGCAGGC CACCGCCGAG	360
	CCGGCGGGCT GTGCTGGGCT CTCCGCGGGT CAAGTGGACT TTCTGTCCC GGGGCGGGGA	420
45	GGCAGAAGTG CTGGTGGCGC GGGGAGTGC CGTCAAGGTG AACGAGGCCT ACCGGTTCCG	480
	CGTGGCACTG CCTGCGTACC CAGCGTCGCT CACCGACGTC TCCCCGGCG CTGAGCGAGC	540
	TGCGCCCCAA CGACTCAGGT ATCTATCGCT GTGAGGTCCA GCACGGCATC GATGACAGCA	600
50	GCGACGCTGT GGAGGTCAAG GTCAAAGGTA TCCCATCCAG ACCCCACGAG AGGCCTGTTA	660
	CGGAGACATG GATGGCTTCC CCGGGGTCCG GAACTATGGT GTGGTGGACC CGGATGACCT	720
55	CTATGATGTG TACTGTATG CTGAAGACCT AAATGGAGAA CTGTTCCCTGG GTGACCCCTCC	780
	AGAGAAGCTG ACATTGGAGG AAGCACGGGC GTACTGCCAG GAGCGGGGTG CAGAGATGTC	840
60	CACCACGGGC CAACTGTATG CAGCCTGGGA TGGTGGCCTG GACCACTGCA GCCCAGGGTG	900

200

	GCTAGCTGAT GGCAGTGTGC GCTACCCCAT CGTCACACCC AGCCAGCGCT GTGGTGGGG	960
	CTTGCCTGGT GTCAAGACTC TCTTCCTCTT CCCCCAACAG ACTGGCTTCC CCAATAAGCA	1020
5	CAGCCGCTTC AACGTCTACT GCTTCGAGA CTCGGCCAG CTTCTGCCAT CCCTGAGGCC	1080
	TCCAACCCAG CCTCCAACCC AGCTTGTGATG GACTAGAGGC TATCGTCACA GTGACAGAGA	1140
	CCCTGGAGGA ACTGCAGCTG CCTCAGGAAG CCACAGAGAG TGAATCCCCTT GGGGCCATCT	1200
10	ACTCCATCCC CATCATGGAG GACGGAGGAG GTGGAGCTC CACTCCAGAA GACCCAGCAG	1260
	AGGGCCCTAG GACGCTCCTA GAATTGAAA CACAATCCAT GGTACCGCCC ACGGGGTTCT	1320
15	CAGAAGAGGA AGGTAAGGCA TTGGAGGAAG AAGAGAAATA TGAAGATGAA GAAGAGAAAG	1380
	AGGAGGAAGA AGAAGAGGAG GAGGTGGAGG ATGAGGCTCT GTGGGCATGG CCCAGCGAGC	1440
	TCAGCAGCCC GGGCCCTGAG GCCTCTCTCC CCACGTGAGCC AGCAGCCCAG GAGGAGTCAC	1500
20	TCTCCAGGC GCCAGCAAGG GCAGTCTGTC AGCCTGGTGC ATCACCACCTT CCTGATGGAG	1560
	AGTCAGAAGC TTCCAGGCCT CCAAGGGTCC ATGGACCACC TACTGAGACT CTGCCACTC	1620
25	CCAGGGAGAG GAACCTAGCA TCCCCATCAC CTTCCACTCT GGTTGAGGCA AGAGAGGTGG	1680
	GGGAGGCAAC TGGTGGTCCT GAGCTATCTG GGTCCTCGA	1720

30

(2) INFORMATION FOR SEQ ID NO: 54:

	(i) SEQUENCE CHARACTERISTICS:	
35	(A) LENGTH: 1117 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:	
	GGCACGAGGC CAAACTTCGG GGGGCTGAGG CGGCGGCCGA GGAGCGCGG ACTCCGGCG	60
	CGGGGAGTCG AGGCATTTCG GCCTGGCTT CGGAGCGTAC CCAGGGCTG AGCCTTTGAA	120
45	GCAGGAGGAG GGGAGGAGAG AGTGGGGCTC CTCTATCGGG ACCCCCTCCC CATGTGGATC	180
	TGCCCCAGGGCG CGGGCGGCCGG AGGAGGGCAG CGAGAAGATG CCCGCCCTGC GCCCCGCTCT	240
50	GCTGTGGCG CTGCTGGCGC TCTGGCTGTG CTGCGCGACC CCCCGCGATG CATTGCACTG	300
	TCGAGATGGC TATGAACCTT GTGTAATGAA AGGAATGTGT GTTACCTACC ACAATGGCAC	360
	AGGATACTGC AAAGGTCCAG AAGGCTCTT GGGGAATAT TGTCAACATC GAGACCCCTG	420
55	TGAGAAAGAAC CGCTGCCAGA ATGGTGGGAC TTGTGTGGCC CAGGCCATGC TGGGGAAAGC	480
	CACGTGCCGA TGTGCCCTCAG GGTTTACAGG AGAGGACTGCA CAGTACTCGA CATCTCATCC	540
60	ATGCTTTGTG TCTCGACCTT GCCTGAATGG CGGCACATGC CATAATGCTCA GCCGGGATAC	600

	CTATGACTGC ACCTGTCAAG TCGGGTTTAC AGGTAAGGAG TGCCAATGGA CCGATGCCCTG	660
5	CCTGTCTCAT CCCTGTGCAA ATGGAAGTAC CTGTACCACT GTGCCAACC ATTTCCCTGCA	720
	AATGCCCTCAC AGGCTTCACA GGGCAGAAGT GTGAGACTGA TGTCAAATGAG TGTGACATTC	780
	CAGGACACTG CCAGCATGGT GGCACCTGCC TCAACCTGCC TGGTTCCCTAC CAGTGCCAGT	840
10	GCCTTCAGGG CTTCACAGGC CAGTACTGTG ACAGCCTGTA TGTGCCCTGT GCACCCCTCGC	900
	CTTGTGTCAA TGGAGGCACC TGTCGGCAGA CTGGTGACTT CACTTTGAG TGCAACTGCC	960
15	TTCCAGAAAC AGTGAGAAGA GGAACAGAGC TCTGGGAAAG AGACAGGGAA GTCTGGAATG	1020
	GAAAAGAAC A CGATGAGAAT TAGACACTGG AAAATATGTA TGTGTGGTTA ATAAAGTGCT	1080
	TTAAAATGAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA	1117

20

## (2) INFORMATION FOR SEQ ID NO: 55:

25	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1903 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:	
	GGCACGAGCT CGGAGAGGCG GCGCCCCCTGA GTACGCCAGG AGCCTCTCTT GCAACTTCTG	60
35	CCACCGCGGG CCACCGCGGC CGCCTGATCC CCCAGAGGAA GGTCGGCGGC GTGGAGCGAT	120
	GACCCCGCGGC GGTCGGGGCG GGCGCCCGGG GCTGCCACAG CCCCCGCCGC TTCTGCTGCT	180
40	GCTGCTGCTG CCGCTGTTGT TAGTCACCGC GGACCCGCCG AAACCTGCCAG GAGTCTACTA	240
	TGCAACTGCA TACTGGATGC CTGCTGAAAA GACAGTACAA GTCAAAATG TAATGGACAA	300
	GAATGGGGAC GCCTATGGCT TTTACAATAA CTCTGTGAAA ACCACAGGCT GGGGCATCCT	360
45	GGAGATCAGA GCTGGCTATG GCTCTAAAC CCTGAGCAAT GAGATCATCA TGTGTTGTGGC	420
	TGGCTTTTGTG GAGGGTTACC TCATTGCCAC ACACATGAAT GACCACTACA CAAACCTCTA	480
50	CCCACAGCTG ATCACGAAAC CTTCCATCAT GGATAAAGTG CAGGATTTTA TGGAGAAGCA	540
	AGATAAGGTG GACCCGGAAA AATATCAAAG AATACAAGAC TGATTCAATT TGGAGACATA	600
	CAGGCTATGT GATGGCACAA ATAGATGGCC TCTATGTAGG AGCAAAGAAG AGGGCTATAT	660
55	TAGAAGGGAC AAAGCCAATG ACCCTGTTCC AGATTCAAGTT CCTGAATAGT GTGGAGATC	720
	TATTGGATCT GATTCCTCA CTCTCTCCCA CAAAAAACGG CAGCCTAAAG GTTTTTAAGA	780
60	GATGGGACAT GGGACATTGC TCCGCTCTTA TCAAGGTTCT TCCCTGGATTT GAGAACATCC	840

	TTTTTGCTCA CTCAAGCTGG TACACGTATG CAGCCATGCT CAGGATATAT AAACACTGGG	900
	ACTTCAACAT CATAGATAAA GATACCAGCA GTAGTCGCCT CTCTTCAGC AGTTACCCAG	960
5	GGTTTTGGA GTCTCTGGAT GATTTTACA TTCTTAGCAG TGGATTGATA TTGCTGCAGA	1020
	CCACAAACAG TGTTGTTAAT AAAACCTGC TAAAGCAGGT AATACCCGAG ACTCTCCGT	1080
10	CCTGGCAAAG AGTCCGTGTGCCAATATGA TGGCAGATAG TGGCAAGAGG TGGCAGACA	1140
	TCTTTCAAA ATACAACCTCT GGCACCTATA ACAATCAATA CATGGTTCTG GACCTGAAGA	1200
	AAGTAAAGCT GAACCACAGT CTTGACAAAG GCACTCTGTA CATTGTGGAG CAAATTCTA	1260
15	CATATGTAGA ATATTCGAA CAAACTGATG TTCTACGGAA AGGATATTGG CCCTCCTACA	1320
	ATGTTCCCTT CCATGAAAAA ATCTACAACG GGAGTGGCTA TCCACTGTTA GTTCAGAACG	1380
	TGGGCTGGA CTACTCTTAT GATTTAGCTC CACGAGCCAA AATTTTCCGG CGTGACCAAG	1440
20	GGAAAGTGAC TGATACGGCA TCCATGAAAT ATATCATGCG ATACAACAAAT TATAAGAAGG	1500
	ATCCTTACAG TAGAGGTGAC CCCTGTAATA CCATCTGCTG CCGTGAGGAC CCTGAACCTCA	1560
25	CCTAACCCAA GTCCCTGGAG GTTGTATGA CACAAAAGGT GGCAGATATY TACCTAGCAT	1620
	CTCAGTACAC ATCCATGCC ATAAGTGGTC CCACAGTACA AGGTGGCTC CCTGTTTTC	1680
	GCTGGACCG TTTCAACAAA ACTCTACATC AGGGCATGCC AGAGGTCTAC AACTTTGATT	1740
30	TTTATTACCAT GAAACCAATT TTGAAACTTG ATATAAAATG AAGGAGGGAG ATGACGGACT	1800
	AGAAGACTGT AAATAAGATA CCAAAGGCAC TATTTTAGCT ATGTTTCTC CATCAGAATT	1860
35	ATGCAATAAA ATATATTAAT TTGTAAAAA AAAAAAAAAA AAA	1903

## 40 (2) INFORMATION FOR SEQ ID NO: 56:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1869 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

50	ACAGCTTTTC GGGGCCGAG TCGCACCCAG CGAAGAGAGC GGGCCGGGA CAAGCTCGAA	60
	CTCCGGCCGC CTCGCCCCTC CCCGGCTCCG CTCCCTCTGC CCCCTCGGGG TCGCGCGCCC	120
	ACGATGCTGC AGGGCCCTGG CTCGCTGCTG CTGCTCTTCC TCGCCTCGCA CTGCTGCCTG	180
55	GGCTCGGCGC CGGGGCTCTT CCTCTTTGGC CAGCCCGACT TCTCCTACAA GCGCANCAAT	240
	TGCAAGCCCA TCCCGGTCAA CCTGCAGCTG TGCCACGGCA TCGAATACCA AACATGCGG	300
60	CTGCCCCAACC TGCTGGGCCA CGAGACCATG AAGGAGGTGC TGGAGCAGGC CGGGCGCTTGG	360

	ATCCCGCTGG TCATGAAGCA GTGCCACCCG GACACCAAGA AGTTCCTGTG CTCGCTCTC	420
5	GCCCCCGCT GCCTCGATGA CCTAGACGAG ACCATCCAGC CATGCCACTC GCTCTGCCTG	480
	CAGGTGAAGG ACCGCTGCAG CCCGGTCATG TCCGCCCTCG GYTTCCCCTG GCCCGACATG	540
	CTTGAGTGCG ACCGTTTCCC CCAGGACAAC GACCTTGCA TCCCCCTCGC TAGCAGCGAC	600
10	CACCTCCCTGC CAGCCACCGA GGAAGCTCCA AAGGTATGTG AAGCCTGCAA AAATAAAAAT	660
	GATGATGACA ACGACATAAT GGAAACGCTT TGTAATAATG ATTTTGCACT GAAAATAAAA	720
15	GTGAAGGAGA TAACCTACAT CAACCGAGAT ACCAAAATCA TCCTGGAGAC CAAGAGCAAG	780
	ACCATTTACA AGCTGAACGG TGTGTCCGAA AGGGACCTGA AGAAATCGGT GCTGTGGCTC	840
	AAAGACAGCT TGCAGTGCAC CTGTGAGGAG ATGAACGACA TCAACGCGCC CTATCTGGTC	900
20	ATGGGACAGA AACAGGGTGG GGAGCTGGTG ATCACCTCGG TGAAGCGGTG GCAGAAGGGG	960
	CAGAGAGAGT TCAAGCGCAT CTCCCGCAGC ATCCGCAAGC TGCAGTGCTA GTCCCCGGCAT	1020
25	CCTGATGGCT CCGACAGGCC TGCTCCAGAG CACGGCTGAC CATTCTGCT CGGGGATCTC	1080
	AGCTCCCGTT CCCCAAGCAC ACTCCTAGCT GCTCCAGTCT CAGCCTGGGC AGCTTCCCCC	1140
	TGCCTTTTGC ACCTTTGCAT CCCCAGCATT TCCTGAGTTA TAAGGCCACA GGAGTGGATA	1200
30	GCTGTTTCA CCTAAAGGAA AAGCCCACCC GAATCTGTGAA GAAATATTCA AACTAATAAA	1260
	ATCATGAATA TTTTATGAA GTTTAAAAAT AGCTCACTTT AAAGCTAGTT TTGAATAGGT	1320
35	GCAACTGTGA CTTGGGTCTG GTTGGTTGTT TTGAGTCAGC TGATTTTCAC	1380
	TTCCCACTGA GGTGTCATA ACATGCAAAT TGCTTCATT TTCTCTGTGG CCCAAACTTG	1440
	TGGGTCACAA ACCCTGTTGA GATAAAGCTG GCTGTTATCT CAACATCTTC ATCAGCTCCA	1500
40	GACTGAGACT CAGTGTCTAA GTCTTACAAC AATTCACTCAT TTTATACCTT CAATGGGAAC	1560
	TTAAACTGTT ACATGTATCA CATTCCAGCT ACAATACTTC CATTATTAG AAGCACATTAA	1620
45	ACCATTCTA TAGCATGATT TCTTCAAGTA AAAGGCAAAA GATATAAATT TTATAATTGA	1680
	CTTGAGTACT TTAAGCCTTG TTTAAAACAT TTCTTACTTA ACTTTTGCAA ATTAAACCCA	1740
	TTGTAGCTTA CCTGTAATAT ACATAGTAGT TTACCTTTAA AAGTTGTAAA AATATTGCTT	1800
50	TAACCAACAC TGTAATATT TCAGATAAAC ATTATATTCT TGTATATAAA CTTTACATCC	1860
	TGTTTTTAC	1869

55

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1259 base pairs

60

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ACCGTGGTCG	TGGCGGGACG	GCGGCTGCAG	CGYGGAGGAG	CTGGGGTCGC	TGTGGGTCGC	60
GAACAGAGCC	CGGGACGTGC	GCGCTTGGTG	CACGATCCTG	AAGGGGAGCT	CCGAGGGGCC	120
10 CGGGTCKCCA	GGGCTGCTGC	GGCCATTCCTC	GGAGCCCCGGC	GGGGGGCCCG	NRAGATACTG	180
GTTTAGGCCG	TCCCAGGGCT	CCGGGCGCAC	CCGKTCGGCG	CTGCTGCAGC	GGAGGGAGCG	240
15 CGGCGGGCGSG	NGGGCTCGGA	GACAGCGTTT	CTCCCGGAAT	CTTCTCGGG	CAGCARGTGG	300
GAAGTGGGAG	CCGGAGCGGC	ACTGGCARCG	TTCTCTCCGC	ANGTCGGCAC	CATGCGCCCT	360
20 GCAGCCCTGC	GGGGGCCCT	GCTGGGCTGC	CTCTGCGCTGG	CGTTGCTTIG	CCTGGGGCGT	420
GCGGACAAGC	GCCTGCGTGA	CAACCATGAG	TGGAAAAAAC	TAATTATGGT	TCAGCACTGG	480
CCTGAGACAG	TATGGGAGAA	AATTCAAAAC	GACTGTAGAG	ACCCCTCGGA	TTACTGGACA	540
25 ATACATGGAC	TATGGCCCGA	TAAAAGTGAA	GGATGTAATA	GATCGTGGCC	CTTCAATTAA	600
GAAGAGATTA	AGGATCTTTT	GCCAGAAATG	AGGGCATACT	GGCCTGACGT	AATTCACTCG	660
30 TTTCCCAATC	GCAGCCGCTT	CTGGAAGCAT	GAGTGGAAA	AGCATGGAC	CTGCGCCGCC	720
CAGGTGGATG	CGCTCAACTC	CCAGAAGAAG	TACTTTGGCA	GAAGCCTGGA	ACTCTACAGG	780
GAGCTGGACC	TCAACAGTGT	GCTTCTAAAA	TTGGGGATAA	AACCATCCAT	CAATTACTAC	840
35 CAAGTTGCAG	ATTTAAAGA	TGCCCTTGCC	AGAGTATATG	GAGTGATACC	CAAAATCCAG	900
TGCCTTCCAC	CAAGCCAGGA	TGAGGAAGTA	CAGACAATTG	GTCAGATAGA	ACTGTGCCTC	960
40 ACTAAGCAAG	ACCAGCAGCT	GCAAAACTGC	ACCGAGCCGG	GGGAGCAGCC	GTCCCCCAAG	1020
CAGGAAGTCT	GGCTGGAAA	TGGGGCCGCC	GAGAGCCGGG	GTCTGAGAGT	CTGTGAAGAT	1080
GGCCCGAGTCT	TCTATCCCCC	ACCTAAAAAG	ACCAAGCATT	GATGCCCAAG	TTTTGGAAAT	1140
45 ATTCTGTTTT	AAAAAGCAAG	AGAAATTCAC	AAACTGCAGC	TTTCTNAAAA	AAAAANAAAA	1200
AAAAAATTGGG	GGGTTTTTTT	GGGGSGCCCG	GGGCCCTTGG	TTTTTCCCCC	CGGGGGGGT	1259

50

(2) INFORMATION FOR SEQ ID NO: 58:

55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1186 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

	CGGCATGGAG AATGGCTCCG CTTCTGTTGC AGCTGGCGGT GCTGGCGCG GCGCTGGCGG	60
5	CCGCAGCCCT CGTACTGATT TCCATCGTTG CATTACAAAC TGCTACAAAA ATGCCAGCAC	120
	TCCATCGACA TGAAGAAGAG AAATTCTTCT TAAATGCCAA AGGCCAGAAA GAAACTTAC	180
	CCAGCATATG GGACTCACCT ACCAAACAAAC TTTCTGTCGT TGTGCCTTCA TACAATGAAG	240
10	AAAAACGGTT GCCTGTGATG ATGGATGAAG CTCTGAGCTA TCTAGAGAAG AGACAGAAC	300
	GAGATCCTGC GTTCACTTAT GAAGTGATAG TAGTTGATGA TGGCAGTAAA GATCAGACCT	360
15	CAAAGGTAGC TTTTAAATAT TGCCAGAAAT ATGGAAGTGA CAAAGTACGT GTGATAACCC	420
	TGGTGAAGAA TCGTGGAAAA GGTGGAGCCA TTAGAATGGG TATATTCACT TCTCGAGGAG	480
	AAAAGATCCT TATGGCAGAT GCTGATGGAG CCACAAAGTT TCCAGATGTT GAGAAATTAG	540
20	AAAAGGGGCT AAATGATCTA CAGCCTTGGC CTAATCAAAT GGCTATAGCA TGTGGATCTC	600
	GAGCTCATTT AGAAAAAGAA TCAATTGCTC AGCGTTCTTA CTTCCGTACT CTTCTCATGT	660
	ATGGGTTCCA CTTTCTGGTG TGGTTCTTT GTGTCAAAGG AATCAGGGAC ACACAGTGTG	720
25	GGTTCAAATT ATTACTCGA GAAGCAGCTT CACGGACGTT TTCATCTCTA CACGTTGAAC	780
	GATGGGCATT TGATGTAGAA CTACTGTACA TAGCACAGTT CTTTAAAATT CCAATAGCAG	840
30	AAATTGCTGT CAACTGGACA GAAATTGAAG GTTCTAAATT AGTTCCATTG TGGAGCTGCC	900
	TACAAATGGG TAAAGACCTA CTTTTTATAC GACTTCGATA TTTGACTGGT GCCTGGAGGC	960
	TTGAGCAAAC TCGGAAAATG AAATTAGGTTG TTGCACTCT TCAGTTGTGT TCTTATGCTT	1020
35	CAGTGTACCA TTTCATTCTA TTGAAACTA AAATTAAAG TAAAGCTGAA ATAAACTTCT	1080
	TGTCACTGTC TGCCTTTGA TAATTTAAA GAAATAACTT TCCATAAGTA AAAAATTATA	1140
40	TATCTCTTG GATATAAAATG ATTTTTAAAAA GATGTTTATT TAAAAA	1186

## 45 (2) INFORMATION FOR SEQ ID NO: 59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 428 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

55	GATCCCCCGG CTGCAGGATT CGGCACGAGT ACTGATTCTT CACTGAGCTT KGTTAGTATA	60
	AGCAGAGTTC CAAGTCTCCC CTAGGGTTGT CTCTACATT CTTTATCATT CCAGTGGGTA	120
60	RGGTTAGCT GGGGAAGGA CATTICATAA GGGTAGTTG GACTGAGCAG TATGGACATT	180

206

TGCTTTTTC ATTACGTACT GTTGTGTTTC CTTGGTAGGT GTGCTTGTT GGTTTAATA TTATTGTGCC AGGGATGGGG AAATGGGGG CGTGTGTGG GAAGAGTACT TATTATGTG 5 TTTTCTTCAG TGAAATTGTT CTTGGTAATT GATACCTCTC TGTTTATTT NTCTCATTCT TTCAAAATAA AACTTTTGA AATTTGAAAA AAAAAAAA AAAAACAATC GGGGGGGGC 10 CCGGTACC	240 300 360 420 428
--	---------------------------------

10

## (2) INFORMATION FOR SEQ ID NO: 60:

15

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 501 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:  GGCACCGAGCT TTCAGCAGGG GACAGCCGA TTGGGGACAA TGGCCCTCT TGCCACATC 25 TTGGTTTCT CTGTGGGTCT CCTCACCATG GCCAAGGCAG AAAGTCCAAA GGAACACGAC CGTTCACTT ACGACTACCA GTCCCTGCAG ATCGGAGGCC TCGTCATCGC CGGGATCCTC 30 TTCATCCTGG GCATCCTCAT CGTGCTGAGC AGAAGATGCC GGTGCAAGTT CAACCAGCAG CAGAGGACTG GGGAAACCCGA TGAAGAGGAG GGAACCTTCC GCAGCTCCAT CGCCCGTCTG TCCACCCGCA GGCGGTAGAA ACACCTGGAG CGATGGAATC CGGCCAGGAC TCCCCTGGCA 35 CCTGACATCT CCCACGGCTCC AACTGCGGCC CCACCGCCCC CTCCGCGCC CCTTCCCCAG CCCTGCCCGGC GCAGACTCCC CCTGCCGCCA AGACTTCCAA TAAAACGTGC GTTCCTCTCG 40 AAAAAAAAAA AAATAAAAAA A	60 120 180 240 300 360 420 480 501
--	--

## 45 (2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1197 base pairs (B) TYPE: nucleic acid 50 (C) STRANDEDNESS: double (D) TOPOLOGY: linear  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:  ACATGATGGN TACCAAAGAA TTCGGCANAG GGCGCGCAGT GCAGCAGGTG CTCATATCG 55 AGTGCCTGCG GGACTTCTG ACGCCCCCGC TGCTGTCCGT GCGCTTCCGG TACGTGGCG CCCCCCAGGC CCTCACCTG AAGCTCCAG TGACCAKCAA CAAGTTCTTC CAGCCCACCG 60	60 120 180
---	------------------

	AGATGGCGGC CCAGGATTTTC TTCCAGCGCT GGAAGCAGCT GAGCCTCCCT CAACAGGAGG	240
	CGCAGAAAAT CTTCAAAGCC AACCACCCCA TGGACCCAGA AGTACTAAG GCCAAGCTTC	300
5	TGGGGTTTGG CTCTGCTCTC CTGGACAATG TGGACCCCAA CCCTGAGAAC TTCGTGGGG	360
	CGGGGATCAT CCAGACTAAA GCCCTGCAGG TGGGCTGTCT CCTTCGGCTG GAGCCCAATG	420
10	CCCAGGCCA GATGTACCGG CTGACCCCTGC GCACCAGCAA GGAGCCCGTC TCCC GTCA CC	480
	TGTGTGAGCT GCTGGCACAG CAGTTCTGAG CCCTGGACTC TGCCCCGGGG GATGTGGCCG	540
	GCACTGGCA GCCCCCTGGGA CTGAGGCAGT TTTGGTGGAT GGGGGACCTC CACTGGTGAC	600
15	AGAGAAAGACA CCACGGTTTG GGGGATGCCT GGGACTTTCC TCCGGCTTT TGTATTTTTA	660
	TTTTTGTCA TCTGCTGCTG TTTACATTCT GGGGGTTAG GGGGAGTCCC CCTCCCTCCC	720
20	TTTCCCCCCC AAGCACAGAG GGGAGAGGGG CCAGGGAAAGT GGATGTCTCC TCCCCTCCC	780
	CCCCACCCCTG TTGTAGCCCC TCCTACCCCC TCCCCATCCA GGGGCTGTGT ATTATTGTGA	840
	GCGAATAAAC AGAGAGACGC TAACAGCCCC ATGTCTGTGT CCATCACCCA CTGTTAGGTA	900
25	GTCAAAGAAG TGGGGTGAGG GCATGCAGAG TGTGGGTGGC CAGNITCGCA GCCCATGGGT	960
	GGGACTCTGG GGAGACAGCA GCAGCAGCAG CCGCCGAAGC CCCAGCTGCA AGGCCACCAG	1020
30	ACGCACCTCT GTGCCTGGTT CCTYAGTCCC CAACACCAGG TAGCAAGCTY TGGGCAGCTG	1080
	GGCCTGGTAG ACCTCATCTT CTGTCCTCTY TGGTGGCCCT GGCTCTGGTG GGAAGTGGGT	1140
	GGAGGTGACC AGGGTATAGA AGTTTCGGAG CTGATTGGAA GAGGATTAAC TTCCCGC	1197
35		

## (2) INFORMATION FOR SEQ ID NO: 62:

40	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 595 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:	
	ATTNANGACK TKYAGCCTYT WATACMATCA TTATAGGGAR AAGCTGGTAC GCCTGMARGT	60
50	ACCGGTCTYGG AATTNCNGGG TCGACCCACG CGTCCGGCAC AGCGGGAGTT GGTTCTGACA	120
	CCAGATGTTC TCTGCTCCTG GTTAATGTCA GTGAGGGCTG GAAGTTGAAT AAATGAGAAC	180
55	AGGAGTGGTC TGGGCCATG TAAATGATCC TCCCTTGAAA GGAGGAACAG CTTTCATCAT	240
	TTGTTCCAGC TAAGCCTTGC ATGCATTATA GATCTGGTGC TAAGCAGTGG GAAAGATCTC	300
	ATAAGTAATG TTTTATGTTC TTTCTGCTC TCCTCTTCTG TWGTTCTTGG CTTGTGGGTT	360
60	GTGTTGTGT GTTAACCTGGA AAATTGCTAT AAGCCAGTTG TCTCTAAGTT TTAAAAACGA	420

ATTAGAAAAA CCATAAAATC TCTGGCCTAT GCACATTGTC CCTGTTTGT GAAAACATTA	480
5 AAGGGTAAAT AAAAAGGAAG GAGAACAGTC AATAATGTGC ATCAAATATA TTCTGAGTTC	540
TAGAGAAATT AATGACCAAG CATTAGAACT AGAAGCAAAA AAAAAAAAAA AAAAA	595

10

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1478 base pairs
  - 15 (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:	
CGCGCGCTGAG GACGCACGGA TGCCTTCCGT GCCTTCCATC AAGATCTCAA TTTTGTGCGC	60
AAGTTCTTAC AGCCCCCTGTT GATTGGAGAG CTGGCTCCGG AAGAACCCAG CCAGGATGGA	120
25 CCCCTGAATG CGCATGGTCG AGGACTTCCG AGCCCTGCAC CAGGCAGCCG AGGACATGAA	180
GCTGTTTGAT GCCAGTCCCA CCTTCTTTCG TTTCTTACTG GGCCACATCC TGGCCATGGA	240
30 GGTGCTGGCC TGGCTCCCTTA TCTACCTCCT GGGTCCTGGC TGGGTGCCA GTGCCCTGGN	300
CGGCCCTTCAT CCTGGCCATC TCTCAGGCTC AGTCCTGGTG TCTGCAGCAT GACCTGGCC	360
ATGCTCCATC TTCAAGAAGW CCTGGTGGAA CCACGTGGCC CAGAAGTTCG TGATGGGCA	420
35 GCTAAAGGCC TTCTCCGCC ACTGGTGGAA CTTCCGCCAC TTCCAGCACC ACGCCAAGCC	480
CAACATCTTC CACAAAGACC CAGACGTGAC GGTGGGCCCG GTCTTCTTCC TGGGGAGTC	540
40 ATCCGTCGAG TATGGCAAGA AGAAACGCAG ATACCTACCC TACAACCAGC AGCACCTGTA	600
CTTCTTCCCTG ATCGGCCCCGC CGCTGCTCAC CCTGGTGAAC TTTGAAGTGG AAAATCTGGC	660
GTACATGCTG GTGTGCATGC AGTGGCCGG A TTTGCTCTGG GCCGCCAGCT TCTATGCCCG	720
45 CTTCTTCTTA TCCTACCTCC CCTTCTACGG CGTCCCTGGG GTGCTGCTCT TCTTGTGTC	780
TGTCAGGGTC CTGGAAAGCC ACTGGTTCGT GTGGATCACA CAGATGAACC ACATCCCCAA	840
50 GGAGATCGGC CACGAGAAGC ACCGGGACTG GGTCACTCT CAGCTGGCAG CCACCTGCAA	900
CGTGGAGCCC TCACTTTCA CCAACTGGTT CAGCGGGCAC CTCAACTTCC AGATCGAGCA	960
CCACCTCTTC CCCAGGATGC CGAGACACAA CTACAGCCGG GTGGCCCCGC TGGTCAAGTC	1020
55 GCTGTGTGCC AACGACAGGCC TCAGCTACGA ATGAAGCCCT TCCTCACCGC GCTGGTGGAC	1080
ATCGTCAGGT CCTGAAGAA GTCTGGTGAC ATCTGGCTGG ACGCCTACCT CCATCAGTGA	1140
60 AGGCAACACC CAGGGGGCA GAGAAGGGCT CAGGGCACCA GCAACCAAGC CAGCCCCGG	1200

	CGGGATCGAT ACCCCCACCC CTCCACTGGC CAGCCTGGGG GTGCCCTGCC TGCCCTCCTG	1260
	GTACTGTGT CTTCCCCCTCG GCCCCCTCAC ATGTGTATTG AGCAGCCCTA TGGCCTGGC	1320
5	TCTGGGCCTG ATGGGACAGG GGTAGAGGGG AGGTGAGCAT AGCACATTTT CCTAGAGCGA	1380
	GAATTGGGGG AAAGCTGTTA TTTTATATT AAAATACATT CAGATGTAAA AAAAAAAA	1440
	AAAAACTCGA GGGGGGGCCC CGGNAACCAA TTCGCCCT	1478
10		

(2) INFORMATION FOR SEQ ID NO: 64:

15	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 2033 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
20	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:	

25	GGCACGAGGA AGAACGCAAA GCTGAGAACAA TGGACGTTAA TATCGCCCCA CTCCGCGCCT	60
	GGGACGATTT CTTCCCGGGT TCCGATCGCT TTGCCCCGGCC GGACTTCAGG GACATTCCA	120
	AATGGAACAA CCGCGTAGTG AGCAACCTGC TCTATTACCA GACCAACTAC CTGGTGGTGG	180
30	CTGCCATGAT GATTTCCATT GTGGGGTTTC TGAGTCCCTT CAACATGATC CTGGGAGGAA	240
	TCGTGGTGGT GCTGGTGGTC ACAGGGTTTG TGTGGGCAGC CCACAATAAA GACGTCCTTC	300
	GCCGGATGAA GAACCGCTAC CCCACGACGT TCGTTATGGT GGTCAATGTTG GCGAGCTATT	360
35	TCCCTATCTC CATGTTTGGG GGAGTCATGG TCTTTGTGTT TGGCATTACT TTTCTTTGC	420
	TGTTGATGTT TATCCATGCA TCGTTGAGAC TTGGAAACCT CAAGAACAAA CTGGAGAATA	480
40	AAATGGAAGG AATAGGTTTG AAGAGGACAC CGATGGCAT TGTCCCTGGAT GCCCTAGAAC	540
	AGCAGGAAGA AGGCATCAAC AGACTCACTG ACTATATCAG CAAAGTGAAG GAATAAACAT	600
	AACTTACCTG AGCTAGGGTT GCAGCAGAAA TTGAGTTGCA GCTTGCCCTT GTCCAGACCT	660
45	ATGTTCTGCT TGGTTTTTG AAACAGGAGG TGCACGTACC ACCCAATTAT CTATGGCAGC	720
	ATGCATGTAT AGGCCGAAC ATTATCAGCT CTGATGTTTC AGAGAGAAGA CCTCAGAAC	780
50	CGAAAGAAAA CCACCACCCCT CCTATTGTGT CTGAAGTTTC ACCTGTGTTT ATGAAATCTA	840
	ATGGGAAATG GATCACACGA TTTCTTTAAG GGAATTAAAA AAAATAAAAG AATTACGGCT	900
	TTTACAGCAA CAATACGATT ATCTTATAGG AAAAAAAAAT CATTGTAAAG TATCAAGACA	960
55	ATACGAGTAA ATGAAAAGGC TGTTAAAGTA GATGACATCA TGTGTTAGCC TGTTCTAAT	1020
	CCCCTAGAAT TGTAATGTGT GGGATATAAA TTAGTTTTA TTATTCTCTT AAAATCAA	1080
60	GATGATCTCT ATCACTTTGC CACCTGTTG ATGTGCAGTG GAAACTGGTT AAGCCAGTTG	1140

	TTCATACTTC CTTTACAAAT ATAAAGATAG CTGTTTAGGA TATTTGTAA CATTGGTAA	1200
5	AATTTTGAA ATGCTAGTAA TGTGTTTCA CCAGCAAGTA TTTGTTGCAA ACTTAATGTC	1260
	ATTTTCCCTA AGATGGTTAC AGCTATGTAA CCTGTATTAT TCTGGACGGA CTTATTAAAA	1320
	TACAAACAGA CAAAAAATAA AACAAAACCTT GAGTTCTATT TACCTTGACAC ATTGTTGTT	1380
10	GTTACAGTGA AAAAATGGT CCAAGAAAAT GTTGCCATT TTGCAATTGT TTCGTTTTA	1440
	ACTGGAACAT TTAGAAAGAA GGAAATGAAT GTGCATTTA TTAATTCCCTT AGGGGCACAA	1500
15	GGAGGACAAT AATAGCTGAT CTTTGAAAT TTGAAAACCG TCTTTAGATG ACCAAGCAAA	1560
	AAGCTTTAAA AAATGGTAAT GAAAATGGAA TGCAGCTACT GCAGCTAATA AAAAATTAA	1620
	GATAGCAATT GTTACAACCA TATGCCCTTA TAGCTAGACA TTAGAATTAT GATAGCATGA	1680
20	GTTTATACAT TCTATTATTT TTCCCTCCCTT TCTCATGTT TTATAAATAG GTAATAAAAAA	1740
	ATGTTTGCC TGCCAATTGA ATGATTTCTGT AGCTGAAGTA GAAACATTAA GGTTCTGTA	1800
25	GCATTAATT GTGAAGACAA CTGGAGTGGT ACTTACTGAA GAAACTCTCT GTATGTCCTA	1860
	GAATAAGAAG CAATGATGTG CTGCTTCTGA TTTTCTTGC ATTGAAATTCTCAGCAAC	1920
	CTACAGCCAT GATCTTACG ACAGTGATAT CACCATGACT TCACAGACAT GGCTAGAACAT	1980
30	CTGTACCCCTT ACCCACATAT GAAGAATAAA ATTGATTAAA GGTTAAAAAA AAA	2033

## 35 (2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 440 base pairs
  - (B) TYPE: nucleic acid
  - 40 (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

45	ATGTTCTTA CTAGAACT GTGTCCAACC TATATAGCCC TAACCTTCCT GGTTTACATT	60
	GTGGCCCTAG TATCTGGCA GCTGTGCATG GAGATAGCCA GAGGAAACAT TTTTCTTCTT	120
50	AATGAATTGG TGACCACATT TTGTTGTTCT TGCCCTCCTAT TATCCGTGCC CTATTGCA	180
	CCTGGTTCT TCTACAGTAG TTTATGTAAA TGTGTTTTG TCCTTGTCT GTCTAGTGA	240
	ATTGGTTCTG TAAACGAAAC CTGGTCTGT AATTCAGTA TATGCTCATA TCTCATCTT	300
55	GGCTCTCCA TTTTCACAGC AGTGATCCCT AAAAGATGTG CCCTAGAGGA TATCCAGAAC	360
	AATCCAATG GATGCTCTCT CGGCTGCACT CCAGGCTGGG AGACAGAGGG AGACTCNATC	420
	TCAAAAAAAA TTAAAAAAA	440
60		

## (2) INFORMATION FOR SEQ ID NO: 66:

5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3301 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

15	GGTCATAAGG GGAGGGTTGN NGTGTGTCCC TCCAGGTTGT GCAGAGGGGA TTAGAAAGTAA	60
	GTAGGTTAGA GGGGAGGTGG AGGGAGTGTG CTGGGGTGTG AGCTTTATG ATGCTGAAAG	120
	GATCATGATA TGCTAAGGAC AGGATAGTGT TGGGTTGTAC ACACAGGTGT AGGCAATCCT	180
20	GGTGGCTAGT ATGTAAGGAGT GAATGTCCTG ACTCCCTTAG AGGGTACCTG NCAGAGTGCC	240
	CTTGGARGGA CTAGTGCTGG AGAAATTAAT AGGAGAGGGG ACGGGCATCC ATTAACCTTT	300
	TCTTGCCCTGC AGCCTGTAGG GTCCAGCGTC AAAGCGAAC ATGGGGTCCA GGGCTGAGCT	360
25	GTGCACTCTC TTAGGCGGAT TCTCCCTTCCT CCTGCTACTG ATACCAGGCG AGGGGGCCAA	420
	GGGTGGATCC CTCAGAGAGA GTCAGGGAGT CTGCTCCAAG CAGACACTGG TGGTCCCGCT	480
30	CCACTACAAC GAGTCCTACA GCCAACCAAGT GTACAAGCCC TACCTGACCT TGTGCGCTGG	540
	GAGCGCATCT GCAGCACTTA CAGGACCATG TACCGCGTTA TGTGGCGGGGA GGTGAGGCCG	600
	GAGGTTCAAGC AGACCCATGC AGTGTGCTGC CAGGGCTGGA AGAAGCGGCA CCCGGGGCG	660
35	CTCACCTGTG AAGCCATCTG CGCCAAGCCT TGCCTGAACG GAGGCCTCTG CGTTAGGCCT	720
	GACCAGTGCG AGTGCCTCCCC CGGCTGGGGGA GGGAAAGCACT GTCATGTGGA CGTGGATGAA	780
40	TGTAGGACCA GCATCACCCCT CTGCTCGCAC CATTGTTTA ATACGGCARG CAGCTTCAMC	840
	TGCGGCTGCC CCATGACCTA GTGCTAGGCG TGGACGGCG CACCTGCATG GAGGGGTCCC	900
	CAGAGCCCCC AACCAAGTGCC AGCATACTCA GCGTGGCCST TCGGGARGCG GAAAAAGATG	960
45	ACGCGCTCTG AAGCAGGAGA TTCACGAGCT GCGAGGCCCT TGAAGCGGCT GGAGCAGTGG	1020
	NCCGGTCAGC TGGGCCCTGG NTCAAGACGGT GCTGCCGTG CGGCCTGAAG WGCTGCAGCC	1080
50	AGAACAGGTG GCTGAGCTGT GGGGCCGGGG TGACCGGATC GAATCTCTCA GCGACCAGGT	1140
	GCTGCTGCTG GAGGAGAGGC TAGGTGCTG CTCCCTGTGAG GACAACAGCC TGGGCCCTGG	1200
	CGTCAATCAT CGATAAGAAG CCTCTACAGC ACCCCTGCCCT CCTAATTAT ACAGAAACCG	1260
55	GACCCACTAA TCCTCTGGGA TTGGCCGACT GTGAGCTGCA GATAAGGCTA TCAGCCACCA	1320
	AAGAGCAATG AACAAATGGAA ACTTCAGAGA GCTGAAGAAA GGGGGAGGCC TGTGTTCTTG	1380
60	GCCTGCCCT GAGTCTTCTG GCTGGGGCA GGTTGCCTGG GCAAGAACTG CTTCTTCAAT	1440

	TCCTTAACAA ATGCAACCAC CAACACCCAG ATCTCTCTCT CTCTTTATTT TCAGTTTTTT	1500
5	TGCTGTATC CAGATAATT AAAAAAACCA ACCACGCAA ACTGGTCCC ACCCTCTCCT	1560
	TTTGCTCCCA GCCTACCTCC CCAGTGTGG GAACAGGTCT GGAGTGAGAG GCAGGGAGTG	1620
	GCTAATGCCN CCAGGAAGAA ATGAAAATG GTCAGAGAG GGGGAAGCCT CAACAGAAAA	1680
10	AGAAATAAAT TAAAAGCCCT CCTATCCCCT CCAGCCAGGG TTCGTTCCCT TCCCCAACTC	1740
	CCCAGGGGGC AGAAGTGAGT GCAGCACCTG ATGCTGCTT CTTCCCCCTTG TGTCTGGTGA	1800
	GATGGTGCAG CAGGGCTGCA GGGGGCTGGG TGGGTCATG TCCACTGAAG AACTGTACTA	1860
15	TGGGGACAGA AAACAGAAA TGTGGAGACT GAACTGGTAT CCCAGAGAGT GCACGACCCT	1920
	GGGCATCTGG GCAAGGGCAG GCATGAGACC TCTGAATTAG AAGGGTCCAG CCCCCACTGA	1980
20	CAGGAGGCTA CACTGGGAGG GAAGGTGAAG GTGCTGAGGA AAGCTCCCAT GATGAGCCTG	2040
	GGAGTGCTTC AGGTATCAGC TTCCAGCCAG AGGGCGAGAA GTCCCTCTCA CAAATGGATG	2100
	AGTCCATTGA ATCCATGGAC TTTGGAGTGG GGGGGATTTG TTCCAAGAA TGGATGAGTC	2160
25	CACTGCCAA TGTGGGTAG AGGGTAGAG AAGACCACAT AGGAAGAGAC TCCACTGGGG	2220
	ATGGAATGTT CCCCTCCCTT GTGTAGGCTG AGTCACTGGA GATGAGGGGG AGGCAACTGT	2280
30	CCCACAGACA ARACAGTAGG AGTGGGGGT CAAGAGTGG AACTGCACCG AGGCAAGAGT	2340
	CCATGGATGG GCCAAGAGG GGGCAGGAGT GGGCTGTAT CCACATTCTCA CTTCAGAACT	2400
	TGAAGATTCC AAAGAGGAGA ATAAGTGGGG AGAGGGAGA CAAGGAAGAG GGTTTKGCC	2460
35	TGCTTCAGGG CCCACTGGGT GGGTAGGTGT GGGGAGGAAG ATGGGACAG ATGGGAGGAG	2520
	AGCTCAGAGC CAGGGTCAC CCACCGCCCC CAGGCTTCTT CAGATAGTCA CCACCACCCC	2580
40	GGCCATCAGT GGAGATTTCC CGAAAACAG TGAAGCATGG AGTGCCTGGAC TCTGTCAGCC	2640
	AGAGCTGGGA CGTCATCTGG TGTCAGCCCT TCCGTGGCA CTGGGGCAG CACCCGCACC	2700
	TGACATTGTC CCGAGGTGAA GCGACGCTCC TTCTTGCACT AGAAGTCTTG GTAGGAGGAC	2760
45	ATGACTATGG GGACAATGGG AACCTGGGCC TGCAC TGCAA GATGGAAGGC GCCACGTTG	2820
	AAGGGCAGCA TGGAGCCATT GTGGTTCTC GTTCCCTCAG GAAACACCCA GACCYTCACG	2880
50	TCCTGGGTGA GCAGGGTCTG GGCGACCTCA GACATGACAC TGATGGCATC CCCCCGGC	2940
	TTCCGGTCGA TGAAGATGAC TCCTGCCAGC CAGCAGGCCA GCGCGCAGAG CCAGCCCACA	3000
	GTANTCGCGC TTGGCAATGG GCACACAGCG GCCTGGCAGT ACCTCCATCA TCCCAAGCAG	3060
55	ATCGAGAGAG CTCTGGTGGT TGGAGACAAC AACATAGGGC TGGGAGGGAG GGAAGTGGTG	3120
	AGCCCCCTCGC ACCTCCACTC GGATCCCGTA CAGGTATTG ATGTGGAGCA GCATTAGACG	3180
60	CAAGATCTTC ATGTTCTCGA CGTTGCGTCC TCGCACGGCA CACACAGGGA TGGCGAGCAC	3240

AGCCAGGAAG AGGATCCAGC CATTGTAGAA GGCCATCTTG AAGAAGTACT TGGCACTGGG 3300

5 G 3301

10 (2) INFORMATION FOR SEQ ID NO: 67:

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1535 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

20	GGCACGAGGT CAAGCGAAAG GATTTCAAGG AACAGATCAT CCACCATGTG TTCACCATCA	60
	TTCTCATCAG CTTTCCTGG TTTGCCAATT ACATCCGAGC TGGGACTCTA ATCATGGCTC	120
	TGCATGGACT CTTCCGATTA CCTGCTGGAG TCAGCCAAGA TGTTTAACTA CGCGGGATGG	180
25	AAGAACACCT GCAACAAACAT CTTCATCGTC TTGCCATTG TTTTTATCAT CACCCGACTG	240
	GTCATCCTGC CCTTCTGGAT CCTGCATTGC ACCCTGGTGT ACCCACTGGA GCTCTATCCT	300
30	GCCTTCTTG GCTATTACTT CTTCAATTCC ATGATGGGAG TTCTACAGCT GTCGCATATC	360
	TTCTGGGCT ACCTCATTTC GGGCATGGCC CACAAGTICA TAACTGGAAA GCTGGTAGAA	420
	GATGAACGCA GTACCGGGAA GAAACAGAGA GCTCAGAGGG GGAGGGAGGCT GCAGCTGGGG	480
35	GAGGAGCAA GAGCCGGCCC CTAGCCAATG GCCACCCAT CCTCAATAAC AACCATCGTA	540
	AGAATGACTG AACCATTATT CCAGCTGCCT CCCAGATTAA TGCATAAAAGC CAAGGAACTA	600
40	CCCCGCTCCC TGGCTATAG GGTCACTTTA AGCTCTGGGG AAAAAGGAGA AAGTGAGAGG	660
	AGAGTTCTCT GCATCCTCCC TCCTTGCTTG TCACCCAGTT GCCTTTAAC CAAATTCTAA	720
	CCAGCCTATC CCCAGGTAGG GGGACGTTGG TTATATTCTG TTAGAGGGGG ACGGTCGTAT	780
45	TTTCCTCCCT ACCCGCCAAG TCATCCTTC TACTGCTTTT GAGGCCCTCC CTCAGCTCTC	840
	TGTGGGTAGG GGTTACAATT CACATTCTT ATTCTGAGAA TTGGCCCA GCTGTTGCC	900
50	TTTGACTCCC TGACCTCCAG AGCCAGGGTT GTGCCTTATT GTCCCATCTG TGGGCCTCAT	960
	TCTGCCAAAG CTGGACCAAG GCTAACCTTT CTAAGCTCCC TAACTGGGC CAGAAACCAA	1020
	AGCTGAGCTT TTAACCTTCT CCCTCTATGA CACAAATGAA TTGAGGGTAG GAGGAGGGTG	1080
55	CACATAACCC TTACCCCTACC TCTGCCAAAA AGTGGGGCT GTACTGGGA CTGCTCGGAT	1140
	GATCTTCTT AGTGCTACTT CTTTCAGCTG TCCCTGTAGC GACAGGTCTA AGATCTGACT	1200
60	GCCTCCTCCT TTCTCTGGCC TCTTCCCCCT TCCCTCTTCT CTTCAAGCTAG GCTAGCTGGT	1260

	TTGGAGTACA ATGGCAACTA ATTCTAAATT TTATTTATTA AATATTTGGG GTTTTGGTTT	1320
	TAAAGCCAGA ATTACGGCTA GCACCTAGCA TTTCAGCAGA GGGACCATT TAGACCAAAA	1380
5	TGTACTGTAA ATGGGTTTTT TTTTAAAATT AAAAGATTAA ATAAAAAATA TTAATAAAAA	1440
	CATGGCAATA AGTGTCAAGAC TATTAGGAAT TGAGAAGGGG GATCAACTAA ATAAACGAAG	1500
10	AGAGTCTTTC TTATGCAAAA AAAAAAAAAA AAAAAA	1535

(2) INFORMATION FOR SEO ID NO: 68:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1244 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

25	GGGCACCCAC CAGCGGCGCC GACCTCAGCG CGCACCTATG GGCTCGCTAC CAGGACATGC GGAGACTGGT GCACGACCTC CTGCCCCCG AGGTCTGCAG TCTCCTGAAC CCAGCAGCCA TCTACGCCAA CAACGAGATC AGCCTGCGTG ACGTTGAGGT CTACGGCTT GACTACGACT	60 120 180
30	ACACCCCTGGC CCAGTATGCA GACGCACTGC ACCCCCAGAT CTTCACTTACCC GCCC GTGACA TCCTGATCGA GCACTACAAG TACCCAGAAG GGATTGCGAA GTATGACTAC AACCCAGCT TTGCCATCCG TGGCCTCCAC TATGACATTC AGAAGAGCCT TCTGATGAAG ATTGACGCCT	240 300 360
35	TCCACTACGT GCAGCTGGGG ACAGCCTACA GGGGCCTCCA GCCTGTGCCA GACGAGGGAGG TGATTGAGCT GTATGGGGGT ACCCAGCACA TCCC ACTATA CCAGATGAGT GGCTTCTATG	420 480
40	GCAAGGGTCC CTCCATTAAG CAGTTCATGG ACATCTTCTC GCTACCGAG ATGGCTCTGC TGTCTGTGT GGTGGACTAC TTTC TGGGCC ACAGCCTGGA GTTGACCAA GCACATCTCT ACAAGGACGT GACGGACGCC ATCCGAGACG TGCATGTGAA GGGCCTCATG TACCA GTGGA	540 600 660
45	TCGAGCAGGA CATGGAGAAG TACATCTGA GAGGGGATGA GACGTTTGCT GTCTGAGCC GCC TGGTGGC CCATGGAAA CAGCTGTTC TCATCACCAA CAGTCCTTTC AGCTTCTAG	720 780
50	ACAAGGGGAT GCGGCACATG GTGGGTCCCG ATTGGCGCCA CTCTTCACTG TGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGCAAG CTTTNCAGAA AACTCGATGA	840 900
55	GAAGGGCTCA CTTCACTGGG ACCGGATCAC CCGCTTGAA AAGGGCAAGA TCTATCGCA GGGAAACCTG TTTGACTTCT TACGCTTGAC GGAATGGCGT GGGCCCCCGCG TGCTCTACTT	960 1020
60	CGGGGACCCAC CTCTATAGTG ATCTGGCGGA TCTCACTGCTG CGGCACGGCT GGCGCACAGG CGCCATCATC CCCGAGCTGG AGCGTGAAGAT CGGCATCATC AACACGGAGC AGTACATGCA	1080 1140

CTCGCTKACG TGGCAGCAGG CGCTCACGGG GCTKCTKGAG CGCATKCAGA CCTATCAGGA	1200
CGCGGAGTTG AGGCAGGTCT TGCTTCCTTG ATGAAAGANC GNNT	1244

5

10 (2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1292 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

GGCACGAGCA GCGACGCCAC TCTGGTGCAG GCGTCTTCT TCCCCCGAG CTGGGCGTGC	60
20 GCGGCCGCAA TGAACCTGGGA GCTGCTGCTG TGGCTGCTGG TGCTGTGCGC GCTGCTCCCTG	120
CTCTTGGTGC AGCTGCTGCG CTTCCCTGAGG GCTGACGGCG ACCTGACGCT ACTATGGGCC	180
25 GAGTGGCAGG GACGACGCCA AGAACATGGAG CTGACTGATA TGGTGGTGTG GGTGACTGGAA	240
GCCTCGAGTG GAATTGGTGA GGAGCTGGCT TACCAAGTTGT CTAAACTAGG AGTTTCTCTT	300
30 GTGCTGTCAAG CCAGAACAGT GCATGAGCTG GAAAGGGTGA AAAGAACATG CCTAGAGAAC	360
GGCAATTAA AAGAAAAAGA TATACTTGTG TTGCCCCCTTG ACCTGACCGA CACTGGTTC	420
CATGAAGCGG CTACCAAAGC TGTTCTCCAG GAGTTTGGTA GAATCGACAT TCTGGTCAAC	480
35 AATGGTGGAA TGTCCCAGCG TTCTCTGTC ATGGATAACCA GCTTGGATGT CTACAGAAAG	540
CTAATAGAGC TTAACTACTT AGGGACGGTG TCCCTGACAA AATGTGTTCT GCCTCACATG	600
40 ATCGAGAGGA AGCAAGGAAA GATTTGTTACT GTGAATAGCA TCCTGGGTAT CATATCTGTA	660
CCTCTTTCCA TTGGATACTG TGCTAGCAAG CATGCTCTCC GGGGTTTTTT TAATGGCCTT	720
CGAACAGAAC TTGCCACATA CCCAGGTATA ATAGTTCTA ACATTTGCC AGGACCTGTG	780
45 CAATCAAATA TTGTGGAGAA TTCCCTAGCT GGAGAACGTCA CAAAGACTAT AGGCAATAAT	840
GGAGACCAGT CCCACAAGAT GACAACCAGT CGTTGTGTG GGCTGATGTT AATCAGCATG	900
50 GCCAATGATT TGAAAGAAGT TTGGATCTCA GAACAACTT TCTTGGTTAG TAACATATT	960
GTGGCAATAC ATGCCAACCT GGGCCTGGTG GATAACCAAC AAGATGGGAA AGAAAAGGAT	1020
TGAGAACATT AAGAGTGGTG TGGATGCAGA CTCTTCTTAT TTAAAGACAAA	1080
55 ACATGACTGA AAAGAGCACC TGTACTTTTC AAGCCACTGG AGGGAGAAAT GGAAAACATG	1140
AAAACAGCAA TCTTCTTATG CTTCTGAATA ATCAAAGACT AATTTGTGAT TTTACTTTTT	1200
60 AATAGATATG ACTTTGCTTC CAACATGGAA TGAAATAAAA AATAAAATAAT AAAAGATTGC	1260

CATGAATCTT GCAAAAAAAA AAAAAAAA AA

1292

5

(2) INFORMATION FOR SEQ ID NO: 70:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1031 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

15	GGGCTGTTGC TTTTGAACAG AACCTATAT TACTCTCCTG GGATCTGAGT TTCTGCAGGT	60
	CATTTGTATG TAGGACCAGG AGTATCTCCT CAGGTGACCA GTTTGCGGA CCCGTATGTG	120
20	GCAAATTCTA AGCTGCCATA TTGAACATCA TCCCACGGG AGTGGTTATG TTGTATCCCC	180
	ATCTTGGCTG GCTTCAGTTT TTGCTGTAGC CCTAGAGCAC TTTGTTGTG GGAGGCTGGC	240
25	CTCTTGCTA CCTCCTTGCA TGGACAGGGG GATGAATATT TACTTTCCA CCTCCTTGCT	300
	TITTCMTCA CTGATACCAC TGAATGGAAC TGGTGCTGTG ACTCCCTGCTG CTGGGGATTT	360
	ATGTCCCGAG ACCTTAGCCT GGCTGAGTGG AGCCTGAGAC CTGCACAAACA GCTCATGGTC	420
30	ATGCATGARA GAGAAGTGGC TGGCCACAGC AGAGGGAAACA GTAACAGCCC AGGGCCTTT	480
	ATTTTGGAA AGGCTGTCGG GGGCTGTTAC TGTCCTCTCT GTTATAAAG CAGACATGTG	540
35	GCCATCTTTT CGCGAGGTA GAGTGGGCTC CTTTCTTTT GGAATCCTTT TCTTCTCCTT	600
	TGGTAGCAGC TCCCTGCCCTC CAGGGCTTCC GCCACCAGCG TCTCTGCTGT GTGCGCAGT	660
	GCAGTGGGGT GCAAGGGCTT TGTTTCTGCC TGCCTGAAAG AGAGGGCTCT GGGGATGGAG	720
40	ATGAGAAACA ACACGCTCTC CTTCAAGACAA TGAGGCATTIC TGCCCTCCTG CTGCCATTCT	780
	TCATCTCCAC TGAGAGCCAG AGCTGGTAGG ACCCGAGTGC CACAGGCATT CTGCATTGCT	840
45	CTACTCTTAG GTTTGIGTGT GTGATCCTTC CCCTCCCTGT CGCCCACTCC TCCCTCCTCT	900
	GGCTATCTA CCCTGCTGTG GGGCTCTTTT ACTACCAGCC TATGCTGTGG GACTGTCATG	960
	GCATTTAGTT CAGAGTGGAN GGGCTTTGGS CTGAAATAAA ATGCAAGTAT TTAAAAAAA	1020
50	AAAAAAAAA A	1031

55 (2) INFORMATION FOR SEQ ID NO: 71:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 855 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

## (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

5	AGCTATTGAC ACTTCCTGGT GGGATCCGAG TGAGGCAGC GGGTAGGGT TGGCGCTCAG	60
	GCGGCGACCA TGGCGTATCA CGGCCTCACT GTGCCTCTCA TTGTGATGAG CGTGTCTGG	120
10	GGCTTCGTCG GCPTCTGGT GCCTGGTT TC ATCCCTAAGG GTCCTAACCG GGGAGTTATC	180
	ATTACCATGT TGGTGACCTG TTCAGTTGC TGCTATCTCT TTTGGCTGAT TGCAATTCTG	240
	GCCCAACTCA ACCCTCTCTT TGGACCGCAA TTGAAAAATG AAACCATCTG GTATCTGAAG	300
15	TATCATTGGC CTTGAGGAAG AAGACATGCT CTACAGTGCT CAGTCTTGA GGTACAGAGA	360
	AGAGAAATGCC TTCTAGATGC AAAATCACCT CCAAACCAGA CCACCTTTCT TGACTTGCC	420
20	GTTTTGGCCA TTACCTGCCT TAAACGTTAA CAGCACATTT GAATGCCCTA TTCTACAATG	480
	CAGCGTGTGTT TCCCTTGCCCT TTTTTGCACT TTGGTGAATT ACGTGCCTCC ATAACCTGAA	540
	CTGTGCCGAC TCCACAAAAC GATTATGTAC TCTTCTGAGA TAGAAGATGC TGTCTCTG	600
25	AGAGATACTG TACTCTCTCC TTGGAATCTG TGGATTTGAA GATGGCTCCT GCCTTCTCAC	660
	GTGGGAATCA GTGAAGTGTGTT TAGAAACTGC TGCAAGACAA ACAAGACTCC AGTGGGTGG	720
	TCAGTAGGAG AGCACGTTCA GAGGGAAGAG CCATCTCAAC AGAATCGCAC CAAACTATAC	780
30	TTTCAGGATG AATTTCTTCT TTCTGCCATC TTTTGGAAATA AATATTTCC TCCCTTCTAW	840
	RRAAAAAAAA ANANN	855
35		

## (2) INFORMATION FOR SEQ ID NO: 72:

## 40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1274 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

	GGCAGAGCTT AGAGTGTGGA AAAGGCAACC AGGTTGGCCG TAAGTGCCTG CTGGAATGCG	60
50	TGTGCCTCCA CACGGGTCTG GGCATCCGGA CTGATAACCA GCCGGCCAGA CTGAGGGATG	120
	GAAGGCAC TG AGATGGGGGC CCGTCCAGGC GGACACCCGC AGAAATGGAG CTTTCTGTGG	180
	TCTCTTGAC TCTGGCTGCC TCTTGCCCTC TCTGTGTCTC TCTTTCTTGG TCTCTCCCTC	240
55	TCTCCTCCCTC AGCCTGGTCT TTCTCTTTGG TGACACACTTA GTTATTGTG TGAGCAATGG	300
	AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GGACACAGCC TAAAAAGGAC	360
60	AAAAAGTTAG AAGACAGCAT AGCAACTCAG CTCAGGGAGC TACCAAGAGAA AAATAGCAAC	420

	TGATGTGGGT GCTTTTTTTT TTTTTTTAAT TTGAATAAAA AGAATTAGAA GTGATGTCCT	480
5	TTTATAAAAT GCCTCTCCC CCTTCCGCC TACAGTCTCT TCCTCTCCCC TTAGAGGGGG	540
	GAAAGTGTAT AACCTACAG GGTTGTGAGT CTGAAAAGAG GATCCCCCTC ACCCCCACCC	600
	TGGGCAGAGC AGTGGGGTT GGGGGTGGG AGAGGGGAC ACAGATCCTG GCACACTGTG	660
10	GATATTCCTT GCAGATTGCA GTCTCTGTG GCCCAAACAG GTTAGGTTAGA CTATGCCCTC	720
	TGGCAGGTGC CACCTTTGG TACCAACATG TTCTGAGGTG TTAGGATTIG GGTTGGGTTT	780
15	TTTTTGTTG TTTTTTTTT CCTTTTGGTC TTTTTTTTTT TCTCCTTTTA AAGAAAAGCT	840
	AAAGGCCGCT GTGAGTCCTG GTGGCAGGCT CTCCATGGAT GTAGCATATC GAAGATAATT	900
	TTTATACTGC ATTTTATGG ATTATTTTGT AATGTGTGAT TCCGTCTGCT GAGGAGGTGG	960
20	GAGGGCTCC AGGGAAAGCC ACCCACCTTC AGTGAGGTG CTCCCCAGCT GAGGCCACCG	1020
	GGCATGGAT GTGGAGGCTG GCGACACACC CTGTGCCTCT CCAAGGCTGG CGCGTGGGG	1080
25	CGTCCAGAGT CTCTCTGGGT CTCAGATGTC CATCTGCCAC CTCTTGTTAA GGCTCTAGCC	1140
	AGAAGGGAGG GTGAGGGTAG AAGAAAGTTA TTCCCGAAGA AAAAAGAAT GAAAAGTCAT	1200
	TGTACTGAAC TGTTTTATA TTTTAAAAG TTACTATTWA AACGTAAAAA AAAGGGGGGG	1260
30	CCCGGTACCC AATT	1274

## 35 (2) INFORMATION FOR SEQ ID NO: 73:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 688 base pairs
  - (B) TYPE: nucleic acid
  - 40 (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

45	GGCACGAGTG GAGGAATGC CAGCTCCAGG ACAGAGGCTC AGGTGCCAA CGGGCAAGGC	60
	AGCCCAGGG GCTGTGCTG TTCAAGTCAG GCTTCCCCGG CCCTCGGCCA CAGCGCTTC	120
50	ACGGGCAGCC CGGGGCCCA CCCCACGCAC TGAAAGAGGCC CCCTGGGTG CCAIGCCCT	180
	GACCTTCCTG CTGGTGCTGC TCACCCCTGGC CACGTCTGCA CACGGCTGCA CAGAAACTTC	240
	CGACGCGGGG AGACATCTA CTGGGGGCC ACAGGGACA GCCAGGACAC AGTGGCTGCT	300
55	GTGCTGAAGC GGAGGCTGCT GCAGCCCTCG CGCCGGGTCA AGCGCTCGCG CGGGAGACCC	360
	CTCTCCCCC CACCCGGAC AGCGGGCCGG AAGGGAGAG CTGGAGTGA CGGCTGGGA	420
	CCTGCCACTG TGGCGTGCAG CTCCCTCCCCG CGCCGCGAGG CGCGACCTC TGCCACGTGG	480
60		

ACCGCGCGCG	GGCGCTCCC	TGGTGGCGAT	GGCGCGCAC	TGGCCGAGCA	CTGCGGGGGC	540
TTTCCTCCCT	GTTGGTTGCT	GAGTGGCGG	CCAAGGGAG	AAAAGGAGCC	GCTTCTGCCT	600
5	CCCTTGCCAA	AACTCCGTTT	CTAATTAAAT	TATTTTTAGT	AGAAAAAAAAA	660
			AAAAAAAAAA	AAAAAAAAAA	AAAAAAA	688

10

(2) INFORMATION FOR SEQ ID NO: 74:

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1890 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

GAGCAGGAGA	GAAGGCCACCG	CCCCACCCCG	CCTCCAAAGC	TAACCCTCGG	GCTTGAGGGG	60	
AAGAGGCTGA	CTGTACGTTTC	CTTCTACTCT	GGCACCACTC	TCCAGGCTGC	CATGGGGCCC	120	
25	AGCACCCCTC	TCCTCATCTT	GTTCCCTTTG	TCATGGTCGG	GACCCCTCCA	AGGACAGCAG	180
	CACCACCTTG	TGGACTACAT	GGAACGCCGA	CTAGCTGCTT	TAGAGGAACG	GCTGGCCCAG	240
30	TGCCAGGACC	AGAGTAGTCG	GCATGCTGCT	GAGCTGCCGG	ACTTCAAGAA	CAAGATGCTG	300
	CCACTGCTGG	AGGTGGCAGA	GAAGGAGCGG	GAGGCACCTCA	GAAC TGAGGC	CGACACCATC	360
35	TCCGGGAGAG	TGGATCGTCT	GGAGCGGGAG	GTAGACTATC	TGGAGACCCA	GAACCCAGCT	420
	CTGCCCTGTG	TAGAGTTGA	TGAGAAGGTG	ACTGGAGGCC	CTGGGACCAA	AGGCAAGGGA	480
	AGAAGGAATG	AGAAAGTACGA	TATGGTGACA	GACTGTCGGCT	ACACAATCTC	TCAAGTGAGA	540
40	TCAATGAAGA	TTCTGAAGCG	ATTTGGTGCC	CCAGCTGGTC	TATGGACCAA	GGATCCACTG	600
	GGGCAAACAG	AGAAGATCTA	CGTGTAGAT	GGGACACAGA	ATGACACAGC	CTTGTCTTC	660
45	CCAAGGCTGC	GTGACTTCAC	CCTTGCCATG	GCTGCCCGGA	AAGCTTCCCG	AGTCCGGGTG	720
	CCCTTCCCT	GGGTAGGCAC	AGGGCAGCTG	GTATATGGTG	GCTTTCTTTA	TTTIGCTCGG	780
	AGGCCTCTG	GAAGACCTGG	TGGAGGTGGT	GAGATGGAGA	ACACTTTGCA	GCTAATCAA	840
50	TTCCACCTGG	CAAACCGAAC	AGTGGTGGAC	AGCTCAGTAT	TCCCAGCAGA	GGGGCTGATC	900
	CCCCCCTACG	GCTTGACAGC	AGACACCTAC	ATCGACCTGG	CAGCTGATGA	GGAAAGGTCTT	960
55	TGGGCTGTCT	ATGCCACCCG	GGAGGATGAC	AGGCACCTGT	GTCTGGCCAA	GTTAGATCCA	1020
	CAGACACTGG	ACACAGAGCA	GCAGTGGGAC	ACACCATGTC	CCAGAGAGAA	TGCTGAGGCT	1080
	GCCTTTGTCA	TCTGTGGAC	CCTCTATGTC	GTCTATAACA	CCCGTCCTGC	CAGTCGGGCC	1140
60	CGCATCCAGT	GCTCCTTGA	TGCCAGCGGA	CCCTGACCCC	TGAACGGGCA	GCACTCCCTT	1200

220

	ATTTTCCCCG CAGATATGGT GCCCATGCCA GCCTCCGCTA TAACCCCCGA GAACGCCAGC	1260
5	TCTATGCCTG GGATGATGGC TACCAGATTG TCTATAAGCT GGAGATGAGG AAGAAAGAGG	1320
	AGGAGGTTTG AGGAGCTAGC CTTGTTTTT GCATCTTCT CACTCCCATA CATTATATT	1380
	ATATCCCCAC TAAATTCTT GTTCCTCATT CTTCAAATGT GGGCCAGTTG TGGCTCAAAT	1440
10	CCTCTATATT TTTAGCCAAT GGCAATCAA TTCTTCAGC TCCTTTGTT CATA CGAAC	1500
	TCCAGATCCT GAGTAATCCT TTTAGAGCCC GAAGAGTCAA AACCCCTCAA GTTCCCTCCT	1560
	GCTCTCCTGC CCCATGTCAA CAAATTCAG GCTAAGGATG CCCCAGACCC AGGGCTCTAA	1620
15	CCTTGTATGC GGGCAGGCC AGGGAGCAGG CAGCAGTGT CTTCCCTCA GAGTGACTTG	1680
	GGGAGGGAGA AATAGGAGGA GACGTCCAGC TCTGTCTCT CTTCCCTCACT CCTCCCTTCA	1740
20	GTGTCCCTGAG GAACAGGACT TTCTCCACAT TGTTTTGTAT TGCAACATT TGCAATTAAA	1800
	GGAAAATCCA CTGCAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAACGG CACGAGGGGG	1860
	GGTCCCGTAC CCAATNGCCC TCACATGCAT	1890
25		

## (2) INFORMATION FOR SEQ ID NO: 75:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1133 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

40	GCCGGTCTGA GTGCCAGAGCT GCTGTCTGCG CGGCCGCTCT GTGGGGCTTC TTTCCCGTCC	60
	TGCTGCTGCT GCTGCTATCG GGGGATGTCC AGAGCTCGGA GGTGCCCGGG GCTGCTGCTG	120
	AGGGATCGGG AGGGAGTGGG GTCCGCATAG GAGATCGCTT CAAGATTGAG GGGCGTGCAG	180
45	TTGTTCCAGG GGTGAAGCCT CAGGACTGGA TCTCGGCCG CCGAGTGTG GTAGACGGAG	240
	AAGAGCACGT CGGTTCCCTT AAGACAGATG GGAGTTTTGT GGTTCATGAT ATACCTCTG	300
	GATCTTATGT AGTGAAGTT GTATCTCCAG CTTACAGATT TGATCCGTT CGAGTGGATA	360
50	TCACTTCGAA AGGAAAATG AGAGCAAGAT ATGTGAATT CATCAAAACA TCAGAGGTTG	420
	TCAGACTGCC CTATCCTCTC CAAATGAAAT CTTCAGGTCC ACCTTCCTAC TTTATTAAAA	480
55	GGGAATCGTG GGGCTGGACA GACTTTCTAA TGAACCCAAT GGTTATGATG ATGGTCTTC	540
	CTTTATTGAT ATTGTGCTT CTGCCTAAAG TGGTCAACAC AAGTGTCTT GACATGAGAC	600
	GGGAAATGGA GCAGTCAATG AATATGCTGA ATTCCAACCA TGAGTTGCCT GATGTTCTG	660
60		

	AGTTCATGAC AAGACTCTTC TCTTCAAAAT CATCTGGCAA ATCTAGCAGC GGCAGCAGTA	720
	AAACAGGCAA AAGTGGGCT GGCAAAAGGA GGTAGTCAGG CCGTCCAGAG CTGGCATTG	780
5	CACAAACACG GCAACACTGG GTGGCATCCA AGTCTTGAA AACCGTGTGA AGCAACTACT	840
	ATAAAACTTGA GTCATCCCGA CGTTGATCTC TTACAACGTGT GTATGTTAAC TTTTAGCAC	900
10	ATGTTTTGTA CTTGGTACAC GAGAAAACCC AGCTTTCATC TTTTGTCTGT ATGAGGTCAA	960
	TATTGATGTC ACTGAATTAA TTACAGTGTC CTATAGAAAA TGCCATTAAT AAATTATATG	1020
	AACTACTATA CATTATGTAT ATTAATTAAA ACATCTTAAT CCAGAAAAAA AAAAAAARAA	1080
15	AACTCGAGGG GGGGCCCGGT ACCCAATTIN CCAAATGGGA GTCGTAAAAA ATC	1133

20 (2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 585 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

30	ATGTTTACAA TGTGTGTAT AAATGGGACA ACTCCTCGCC CTCTACCTGT CCCCTCCCCC	60
	TTTGGTTGTA TGATTTCTT CTTTTTTAAG AACCCCTGGA AGCAGCGCCT CCTTCAGGGT	120
	TGGCTGGGAG CTCGCCCAT CCACCTCTTG GGGTACCTGC CTCTCTCTCT CCTGTGGTGT	180
35	CCCTTCCCTC TCCCAGTGC TCGGTGTTCA GTGGGTATA TTTCTCTCC CAGACATGGG	240
	GCACACGCC CAAGGGACAT GATCCTCTCC TTAGTCCTAG CTACATGGGC TCTTTATAAG	300
40	GAGTTGGGG GTAGAGGCAG GAAATGGGAA CCGAGCTGAA GCAGAGGCTG AGTTAGGGGG	360
	CTAGAGGACA GTGCTCCTGG CCACCCAGCC TCTGCTGAGA ACCATTCCTG GGATTAGAGC	420
	TGCCCTTCCC AGGGAAAAAG TGTGCTCTCC CCGACCCCTCC CGTGGGCCCT GTGGGTGTGAT	480
45	GCTGTGCTG TATATTCTAT ACAAAAGGTAC TTGTCCTTTC CCTTTGTAAA CTACATTGAA	540
	CATGGATTA ACCAGTATAA ACAGTTAAAA AAAAAAAA AAAAA	585

50

(2) INFORMATION FOR SEQ ID NO: 77:

55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 577 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

	GGCACGAGGC CTTGCAGAAC TTCTACTTGC CTGCCCTCCCT GCCTCTGGCC ATGGCCTGCC	60
5	GGTGCTTCAG CTTCCCTCTG ATGGGGACCT TCCTGTCA GTTCCCAGACA GTCCTGGCCC	120
	AGCTGGATGC ACTGCTGGTC TTCCCAGGCC AAGTGGCTCA ACTCTCTGC ACAGCTCAGCC	180
10	CCCAGCACGT CACCATCAGG GACTACGGTG TGTCTGGTA CCAGCAGCGG GCAGGCAGTG	240
	CCCCTCGATA TCTCCTCTAC TACCGCTCGG AGGAGGATCA CCACCGGCCG GCTGACATCC	300
	CCGATCGATT CTCGGCAGCC AAGGATGAGG CCCACAATGC CTGTCCTCTC ACCATTAGTC	360
15	CCGTGCAGCC TGAAGACGAC GCGGAACTACT ACAGCTCTGT TGGCTACGGC TTTAGTCCCT	420
	AGGGGTGGGG TGTGAGATGG GTGCCTCCCC TCTGCCTCCC ATTTCTGCCCT CTGACCTTGG	480
	GTCCTTTTA AACCTTCCTCT GAGCCTTGCT TCCCCTCTGT AAAATGGTTT AATAATATTC	540
20	AACATGTCAA CAACAAAAAA NAAAAAWAAA AACTCGA	577

25

## (2) INFORMATION FOR SEQ ID NO: 78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2278 base pairs
- 30 (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

35	GTAATTCCGGC ACGAGGCGCC CAACATGGCG GGTGGGGCT GGGGCCCGCA SCTAACGGCG	60
	CTCCTGGCCG CCTGGATCGC GGCTGTGGCG GCGACGGCAG GCCCGAGGA GGCGCGCTG	120
40	CCGCGGGAGC AGAGCCGGGT CCAGCCCCATG ACCGCTCCA ACTGGACGCT GGTGATGGAG	180
	GGCGAGTGGA TGCTGAAATT TTACGCCCCA TGGTGTCCAT CCTGCAGCA GACTGATTCA	240
	GAATGGGAGG CTTTGCAAA GAATGGTGAA ATACTTCAGA TCAGTGTGGG GAAGGTAGAT	300
45	GTCATTCAAG AACCAAGTTT GAGTGGCCGC TTCTTTGTCA CCACCTCTCC AGCAATTTTT	360
	CATGCAAAGG ATGGGATATT CGCCGTTAT CGTGGCCAG GAATCTCGA AGACCTGCAG	420
50	AATTATATCT TAGAGAAGAA ATGGCAATCA GTCGAGCCTC TGACTGGCTG GAAATCCCCG	480
	GCTTCTCTAA CGATGTCTGG AATGGCTGGT CTTTTAGCA TCTCTGGCAA GATATGGCAT	540
	CTTCACAACT ATTTCACAGT GACTCTTGGA ATTCCCTGCCTT GGTGGTCTTA TGTCTTTTC	600
55	GTCATAGCCA CCTGGTTTT TGGCCTTTT ATGGGCTGG TCTTGGGGT AATATCAGAA	660
	TGTTTCTATG TGCCACTTCC AAGGCATTTA TCTGAGGGTT CTGAGCAGAA TCGGAGATCA	720
60	GAGGAGGCTC ATAGAGCTGA ACAGTTGCAG GATGCGGAGG AGGAAAAAGA TGATTCAAAT	780

	GAAGAAGAAA ACAAAAGACAG CCTTGTAGAT GATGAAGAAG AGAAAAGAAGA TCTTGGCGAT	840
5	GAGGATGAAG CAGAGGAAGA AGAGGGAGGAG GACAACTTGG CTGCTGGTGT GGATGAGGAG	900
	AGAAAGTGAGG CCAATGATCA GGGGCCCCCA GGAGAGGACG GTGTGACCCG GGAGGNAAGT	960
10	AGAGCCTGAG GAGGCTGAAG AAGGCATCTC TGAGCAACCC TGCCCAGCTG ACACAGAGGT	1020
	GGTGGAAAGAC TCCTTGAGGC AGCGTAAAAG TCAGCATGCT GNCAAGGGAC TGTAGATTAA	1080
	ATGATGCGTT TTCAAGAATA CACACAAAAA CAATATGTCA GCTTCCCTTT GGCCTGCAGT	1140
15	TTGTACCAAA TCCTTAATTTC TTCTCTGAATG AGCAAGCTTC TCTTAAAAGA TGCTCTCTAG	1200
	TCATTTGGTC TCATGGCAGT AAGCCTCATG TATACTAAGG AGAGTCTTCC AGGTGTGACA	1260
	ATCAGGATAT AGAAAAAACAA ACGTAGTGTN TGGGATCTGT TTGGAGACTG GGATGGGAAC	1320
20	AAGTTCATTT ACTTACGGGT CAGAGAGTCT CGACCAGAGG AGGCCATTCC CAGTCCTAAT	1380
	CAGCACCTTC CAGAGACAAG GCTGCAGGCC CTGTGAAATG AAAGCCAAGC AGGAGCCTTG	1440
	GNTCTGAGGC ATCCCCAAAG TGTAACGTAG AAGCCTTGCA TCCTTTCTT GTGTAAAGTA	1500
25	TTTATTTTG TCAAATTGCA GGAAACATCA GGCACCACAG TGCATGAAAA ATCTTTCACA	1560
	GCTAGAAATT GAAAGGGCCT TGGGTATAGA GAGCAGCTCA GAAGTCATCC CAGCCCTCTG	1620
30	AATCTCCTGT GCTATGTTTT ATTTCTTACC TTTAATTTTT CCAGCATTTC CACCATGGC	1680
	ATTCAAGGCTC TCCACACTCT TCACTATTAT CTCTTGGTCA GAGGACTCCA ATAACAGCCA	1740
	GGTTTACATG AACTGTGTTT GTTCATTCTG ACCTAAGGG TTTAGATAAT CAGTAACCAT	1800
35	AACCCCTGAA CCTGTGACTG CCAAACATCT CAAATGAAAT GTTGTGGCCA TCAGAGACTC	1860
	AAAAGGAAGT AAGGATTTA CAAGACAGAT TAAAAAAAAA TTGTTTGTC CAAAATATAG	1920
40	TTGTTGTGA TTTTTTTTTA AGTTTCTAA GCAATATTTT TCAAGCCAGA AGTCCTCTAA	1980
	GTCTTGCCAG TACAAGGTAG TCTTGTGAAG AAAAGTTGAA TACTGTTTG TTTTCATCTC	2040
	AAGGGGTTCCTGGT AACTACTTTA ATAATAACTA AAAAACCACT TCTGATTTTC	2100
45	CTTCAGTGAT GTGCTTTGG TGAAAGAATT AATGAACCTCC AGTACCTGAA AGTGAAGAT	2160
	TTGATTTGT TTCCATCTTC TGTAATCTTC CAAAGAATTAA TATCTTGTA AATCTCTCAA	2220
50	TACTCAATCT ACTGTAAAGTA CCCAGGGAGG CTAATTCTYT TAAAAAAAAA AAAAAAAA	2278

55 (2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1143 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

60

## (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

5	CCCCCTCCAAC TCTCAACCCA CTTCTCCAGC CAGCGCCCCA GCCCTCCCGC CGCCCGCTCG	60
	CAGGTCCCGA GGAGCCAGA CTGTGTCCT GACAATGGGA ACAGCCGACA GTGATGAGAT	120
10	GGCCCCGGAG GCCCCACAGC ACACCCACAT CGATGTGCAC ATCCACCAGG AGTCTGCCCT	180
	GGCCAAGCTC CTGCTCACCT GCTGCTCTGC GCTGCGGCC CGGGCCACCC AGGCCAGGG	240
	CAGCAGCCGG CTGCTGGTGG CCTCGTGGGT GATGCAGATC GTGCTGGGA TCTTGAGTGC	300
15	AGTCCTAGGA GGATTTTTCT ACATCCCGA CTACACCCCTC CTCGTCACCT CGGGAGCTGC	360
	CATCTGGACA GGGGCTGTGG CTGTGCTGGC TGGAGCTGCT GCCTTCATTT ACGAGAAACG	420
	GGGTGGTACA TACTGGGCCG TGCTGAGGAC TCTGCTAGCG CTGGCAGCTT TCTCCACAGC	480
20	CATCGCTGCC CTCAAACATT GGAATGAAGA TTTCCGATAT GGCTACTCTT ATTACAACAG	540
	TGCCCTGCCGC ATCTCCAGCT CGAGTGACTG GAACACTCCA GCCCCCACTC AGAGTCCAGA	600
25	AGAACAGTCAG AGGCTACACC TATGTACCTC CTTCATGGAC ATGCTGAAGG CCTTGTTCAG	660
	AACCCCTCAG GCCATGCTCT TGCGGTCTG GATTCTGCTG CTTCTGGCAT CTCTGGCCCC	720
	TCTGTGGCTG TACTGCTGGA GAATGTTCCC AACCAAAGGG AAAAGAGACC AGAAGGAAAT	780
30	GTTGGAAAGTG AGTGGAATCT AGCCATGCCT CTCCGTGATTA TTAGTGCCTG GTGCTCTGC	840
	ACCGGGCGTC CCTGCATCTG ACTGCTGGAA GAAGAACAG ACTGAGGAAA AGAGGCTCTT	900
35	CAACAGCCCC AGTTATCCTG GCCCCATGAC CGTGGCCACA GCCCTGCTCC AGCAGCACTT	960
	GCCCCATTCCCT TACACCCCTT CCCCATCCTG CTCCGCTTCA TGCCCCCTCC TGAGTAGTCA	1020
	TGTGATAATA AACTCTCATG TTATTGTTNN NAAAAAAA AAAAAAAA AATTGGGGG	1080
40	GGGGCCCGTA CCCATTGGGC CTNNGGGGN GGTTTAAAT TAATGGGGG GGTTTAAAG	1140
	GGN	1143

45

## (2) INFORMATION FOR SEQ ID NO: 80:

## 50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 557 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

GGCAGAGAGC AGATGGCCTT GACACCAGCA GGGTGACATC CGCTATTGCT ACTTCTCTGC	60
60 TCCCCCACAG TTCTCTGGA CTTCTCTGGA CCACAGTCCT CTGCCAGACC CCTGCCAGAC	120

	CCCAGTCCAC CATGATCCAT CTGGGTACACA TCCTCTTCCT GCTTTGCTC CCAGTGGCTG	180
5	CAGCTCAGAC GACTCCAGGA GAGAGATCAT CACTCCCTGC CTTTACCCCT GGCACITCAG	240
	GCTCTTGTTC CGGATGTGGG TCCCTCTCTC TGCCGCTCCT GGCAGGCCCTC GTGGCTGCTG	300
10	ATGCGGTGGC ATCGCTGCTC ATCGTGGGG CGGTGTTCTC GTGCCACGC CCACGCCGCA	360
	CCCCCGCCA AGAAGATGGC AAAGTCTACA TCAACATGCC AGGCAGGGC TGACCCCTCCT	420
15	GCAGCTTGGG CCTTTGACTT CTGACCCCTCT CATCCTGGAT GGTGTGTGGT GGCACAGGAA	480
	CCCCCGCCCC AACTTTGGG TTGTAATAAA ACAATTGAAA CACCAAAAAA AAAAAAAAAA	540
	AAAAAAAAAA AANTCGA	557

20 (2) INFORMATION FOR SEQ ID NO: 81:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 795 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

30	GCCGGGGCGA TGTGGAGCGC GGGCCGCGGC GGGGCTGCCT GGCCGGTGCT GTGGGGCTG	60
	CTGCTGGCGC TGTAGTGCC GGGGGTGGT GCCGCCAAGA CCGGTGCCGA CTCGTGACCT	120
35	GCGGGTCGGT GCTGAAGCTG CTCAAATACGC ACCACCCGGT GCGCTGCACT CCCACGACAT	180
	CAAATACGGA TCCGGCAGCG GCCAGCAATC GGTGACCGGC GTAGAGGGGT CGGACGACGC	240
40	MAATAGCTAC TGGGGGATCC GCGGGGGCTC GGAGGGGGG TGCCTGGGG GGTCCCCGGT	300
	GCGCTGGGG CAGGCGGTGA GGTCACGCA TGTCTTACG GGCAAGAACY TGCACACGCA	360
	CCAYTTCCCG TCGCCGCTGT CCAACAACCA GGAGGTGAGT GCCTTTGGGG AAGACGGCGA	420
45	GGCGACGAC CTGGACCTAT GGACAGTGCG CTGCTCTGGA CAGCACTGGG AGCGTGAGGC	480
	TGCTGTCCCT TCCAGCATGT GGGCACCTCT GTGTTCTGT CAGTCACGGG TGAGCAGTAT	540
50	GGAAGCCCCA TCCGTGGCA GCATGAGGTC CACGGCATGC CCAGTGCCTA CACGCACAAT	600
	ACGTGGAAGG CCATGGAAGG CATCTTCATC AAGCCTAGTG TGGAGCCCTC TGCAGGTAC	660
	GATGAACCTCT GAGTGTGTGG ATGGATGGGT GGATGGAGGG TGGCAGGTGG GGCGTCTGCA	720
55	GGGCCACTCT TGGCAGAGAC TTTGGGTTTG TAGGGCTCT CAAGTGCCTT TNIGATTAAA	780
	GAATGTGGT CTATG	795

## (2) INFORMATION FOR SEQ ID NO: 82:

	(i) SEQUENCE CHARACTERISTICS:	
5	(A) LENGTH: 1324 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:	
	NAGGCTTTAA AGCGCCTACC CTGCCTGCAG GTGAGCAGTG GTGTGTGAGA GCCAGGGTC	60
	CCTCTGCCTG CCCACTCAGT GGCAACACCC GGGACCTGTT TTGTCCTTIG TGGAGCCTCA	120
15	GCAGTTCCCT CTTTCAGAAC TCACTGCCAA GAGCCCTGAA CAGGAGCCAC CATGCAGTCC	180
	TTCAGCTTCA TTAAGACCAT GATGATCCTC TTCAATTTC TCATCTTCT GTGTGGTGCA	240
20	GCCCTGTTGG CAGTGGGCAT CTGGGTGTCA ATCGATGGGG CATCCTTCT GAAGATCTTC	300
	GGGCCACTGT CGTCCAGTGC CATGCAGTT GTCAACGTGG GCTACTTCCT CATGCCAGCC	360
	GGCGTTGTGG TCTTTGCTCT TGTTTCCCTG GGCTGCTATG GTGCTAAGAC TGAGAGCAAG	420
25	TGTGCCCTCG TGACGTTCTT CTTCATCCTC CTCCCTCATCT TCATTGCTGA GGTTGCAGCT	480
	GCTGTGGTCG CCTTGGTGTA CACCACAATG GCTGAGCACT TCCGTACGTT GCTGGTAGTG	540
30	CCTGCCATCA AGAAAGATTA TGGTCCCAG GAAGACTTCA CTCAAGTGTG GAACACNACC	600
	ATGAAAGGGC TCAAGTCTG TGGCTTCACC AACTATAACGG ATTTCAGGAA CTCACCCCTAC	660
	TTCAAAGAGA ACAGTGCCTT TCCCCATTC TGTTGCAATG ACAACGTCAC CAACACAGCC	720
35	AATGAAACCT GCACCAAGCA AAAGGCTCAC GACCAAAAAG TAGAGGGTG CTTCAATCAG	780
	CTTTTGATG ACATCCGAAC TAATGCAGTC ACCGTGGGTG GTGTGGCAGC TGGATTGGG	840
40	GGCCTCGAGC TGGCTGCCAT GATTTGKTCATG ATGTATCTGT ACTGCAATCT ACAATAAGTC	900
	CACTTCTGCC TCTGCCACTA CTGCTGCCAC ATGGGAACCTG TGAAGAGGCA CCCTGGCAAG	960
	CAGCAGTGAT TGGGGGAGGG GACAGGATCT AACAAATGTC CTTGGGCCAG AATGGACCTG	1020
45	CCCTTTCTGC TCCAGACTTG GGGCTAGATA GGGACCACTC CTTTTAGCGA TGCCTGACTT	1080
	TCTTCCATT GGTGGGTGGA TGGGTGGGG GCATTCCAGA GCCTCTAAGG TAGCCAGTTC	1140
50	TGTTGCCAT TCCCCAGTC TATTAAACCC TTGATATGCC CCCTAGGCCT AGTGGTGATC	1200
	CCAGTGCTCT ACTGGGGGAT GAGAGAAAGG CATTTCATAG CCTGGGCATA AGTGAAATCA	1260
	GCAGAGCCTC TGGGTGGATG TGTAGAAGGC ACTTCAAAT GCATAAACCT GTTACAATGT	1320
55	TTAAA	1324

## (2) INFORMATION FOR SEQ ID NO: 83:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1494 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

10	CTCAGGCTTC TGTCTCACTT TTCCGGGGGG GGGATTAGGG CAAGGAGGGC ATGAGGGACT	60
	GTCTCTCCCT AAAACCCAGA CCCCTGTTCC CCACTCAGTT CTTCTTCATC CTCCTCCTCA	120
15	TCTTCATTGC TGAGGTTGCA GCTGCTGTGG TCGCCTTGGT GTACACCACA ATGGTGAGAC	180
	ACTGGGATGG AGGAAGGGAA GAAGATTGGG CAAAACCTG GGAGTGGCT GTGGCCTGTG	240
20	AATGGCCACC TTCTGTACCA GCCCTAAAC ACTGGCCTGC CTCACCCAGG CTGAGCACTT	300
	CCTGACGTTG CTGGTAGTGC CTGCCATCAA GAAAGATTAT GGTTCCCAGG AAGACTTCAC	360
	TCAAGTGTGG AACACCACCA TGAAAGGGT AAGGTTGGCT GGGGGAGGTT TTAGGGTGGA	420
25	GAGAAAGAAG CAAGGCCCCA CCTCCACCCCT CATCTTGTCT CCAGCTCAAG TGCTGTGGCT	480
	TCACCAACTA TACGGATTTT GAGGACTCAC CCTACTTCAA AGAGAACAGT GCCTTTCCCC	540
30	CATTCTGTTG CAATGACAAC GTCACCCAAC ACAGCCCAAT GAAACCTGCA CCAAGAAAA	600
	GGCTCACSAC CNAAAARTAN AGGTGTGGC TGGCATGAGT GGGTGGGGAC TGTTTTCAIG	660
	GCCTCAGAGT GGCAAACGGG GATGGGAGTA GGGCAGCTGC CAACTATAAA TGCTTTTTC	720
35	TCTTCCYGAAG GGGTTGCTTC AATCAGCTTT TGTATGACAT CCGAACTAAT GCAGTCACCG	780
	TGGGTGGTGT GGCACTGGA ATTGGGGGCC TCGAGGTAAG CAGATSAGGA GCTGGGACTG	840
	GGACATGGGC ATGAGACCAAG GGCTGCTCAA CCCATCTGAG GCCTCTCTGG AGGAAACAGA	900
40	CTTCTAACTG GGCCCTCAGGT AGGGTGTCTG TGGGACAGGC TTCAGGATCC CTATCATGTT	960
	CCCTCATCTC TCCCCTGTTCC TCCCCTCTCCA GCTGGCTGCC ATGATTGTGT CCATGTATCT	1020
45	GTACTGCAAT CTACAATAAG TCCACTTCTG CCTCTGCCAC TACTGCTGCC ACATGGGAAC	1080
	TGTGAAGAGG CACCTGGCA AGCAGCAGTG ATTGGGGAG GGGACAGGAT CTAACAATGT	1140
	CACTTGGGCC AGAATGGACC TGCCCTTTCT GCTCCAGACT TGGGGCTAGA TAGGGACCAC	1200
50	TCCTTTTAGC GATGCCCTGAC TTCCCTCTCCA TTGGTGGGTG GATGGGTGGG GGGCATTCCA	1260
	GAGCCTCTAA GGTAGCCAGT TCTGTTGCCCT ATTCCCCAG TCTATTAAAC CCTTGATATG	1320
55	CCCCCTAGGC CTAGTGGTGA TCCCAGTGCT CTACTGGGG ATGAGAGAAA GGCATTTAT	1380
	AGCCTGGCA TAAGTGAAT CAGCAGAGCC TCTGGGTGGA TGTGTAGAAG GCACATTCAA	1440
60	ATGCATAAAC CTGTTACAAT GTTAAAAAAA AAAAAAAA AACTCGACTC TGCC	1494

## (2) INFORMATION FOR SEQ ID NO: 84:

5

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1285 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

	GCTACGTGGC TGGCATGCAT GGGAACGAGG CCCTGGGGCG GGAGTTGCTT CTGCTCCTGA	60
15	TGCAGTTCT GTGCCATGAG TTCCTGCGAG SGAACCCACG GGTGACCCGG CTGCTCTCTG	120
	AGATGCGCAT TCACCTGCTG CCCTCCATGA ACCCTGATGG CTATGAGATC GCCTACCACC	180
20	GGGGTTCAGA RCTGGTGGGC TGGGCCGARG GCCGCTGGAA CAACCAGAGC ATCGATCTTA	240
	ACCATAATTT TGCTGAMCTC AACACACCCAC TGTGGGAAGC ACAGGACGAT GGGAGGTGC	300
25	CCCACATCGT CCCCCAACCAT CACCTGCCAT TGCCCCACTTA CTACACCCCTG CCCAATGCCA	360
	CCGTGGCTCC TGAAACGCGG GCAGTAATCA AGTGGATGAA GCGGATCCCC TTTGTGCTAA	420
	GTGCCAACCT CCACGGGGGT GAGCTCGTGG TGTCCCTACCC ATTGACATG ACTCGCACCC	480
30	CGTGGGCTGC CCGCGAGCTC ACGCCCACAC CAGATGATGC TGTGTTTCGC TGGCTCAGCA	540
	CTGTCTATGC TGGCAGTAAT CTGCCATGC AGGACACCAAG CCGCCGACCC TGCCACAGCC	600
	AGGACTTCTC CGTGACCGC AACATCATCA ACAGGGCYTG ACTNGGCACA CGGTCCCCGG	660
35	GANGCATGAA TGAYTTCAAC TACCTACACA CCAACTGCTT TGAGGTCACT GTGGAGCTGT	720
	SCTGTGACAA GTTCCCTCAC GAGAATGAAT TGCCCCAGGA GTGGGAGAAC AACAAAGACG	780
40	CCCTCCCTCAC CTACCTGGAG CAGGTGCCA TGGCATTGAG AGGAGTGGTG AGGGACAAGG	840
	ACACGGAGCT TGGGATTGCT GACGCTGTCA TTGCGTGGGA TGGGATTAAC CATGACGTGA	900
45	CCACGGCGTG GGGGGGGAT TATTGGCGTC TGCTGACCCC AGGGGACTAC ATGGTGACTG	960
	CCAGTKCCGA GGGCTACCAT TCAGTGACAC GGAACGTGCG GGTACACCTTT GAAGAGGGCC	1020
	CCTTCCCCCTG CAATTTCGTG CTCACCAAGA CTCCCAAACA GAGGCTGCCG GAGCTGCTGG	1080
50	CAGCTGGGGC CAAGGTGCC CCGGACCTTC GCAGGGCCT GGAGCGGCTA AGGGGACAGA	1140
	AGGATTGATA CCTGCGTTT AAGAGCCCTA GGGCAGGCTG GACCTGTCAA GACGGGAAGG	1200
	GGAAGAGTAG AGAGGGAGGG ACAAAAGTGAG GAAAAGGTGC TCATTAAGC TACCGGGCAC	1260
55	CTTAAAAAAA AAAAAAAAAN AAAAAA	1285

## (2) INFORMATION FOR SEQ ID NO: 85:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 394 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

10	GCGCGCTCTA GGAACTAGTG GATCCCCGG GNCTGCAGGT GTGGAGTGGG CCATCGTAAA	60
	TAGTATCTGT GCATAAGGTG GTTGTGCGAT AAATGAGTTA ATGTATGCAA AGCCCTTGGC	120
15	CCAGAGCCGG CCCAGAGCAT TGTGTAAGTS CTGGCAGGCG TCATGATGGA GATATCATGT	180
	CTCCCTCTRT TGATTCAAGGA TTCTGATGAG ATGGAGGATG GGCCTGGGT TCAGGATTAG	240
20	GCCTTGAGGC ACTGCTCCAG CCTCCTTGT GGGCCTGTC ACCCTGGCT TCATCGGGCC	300
	GTARCAAGTC TCCCCTCTCC CACTYTGCAG CAGARGTGT CAAGAACTGC CTGCTCACGG	360
	TTCGTGTTCT GCAAGGCCAT CGCCTAACCT CTAA	394

25

## (2) INFORMATION FOR SEQ ID NO: 86:

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1925 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:
----	---

	AGTGAAGGGA GCTGGCCGTG CGACTGGGCT TCGGGCCCTG TGCCAGAGGA GCANGCCTTC	60
40	CTGAGCAGGA GGAAGCAGGT GGIGGCCGCG GCCTTGAGGC AGGCCCTGCA GCTGGATGGA	120
	GACCTGCAGG AGGATGAGAT CCCAGTGGTA GCTATTATGG CCACTGGTGG TGGGATCCGG	180
45	GCAATGACTT CCCTGTATGG GCAGCTGGCT GGCTGAAGG AGCTGGCCT CTTGGATTGC	240
	KTCTCCTACA TCACCGGGGC CTCGGCTCC ACCTGGCCT TGGCCAACCT TTATAAGGAC	300
	CCAGAGTGGT CTCAGAAGGA CCTGGCAGGG CCCACTGAGT TGCTGAAGAC CCAGGTGACC	360
50	AAGAACAAAGC TGGGTGTGCT GGCCCCCAGC CAGCTGCAGC GGTACCGGCA GGAGCTGGCC	420
	GAGCGTGCC CTTGGGCTA CCCAAGCTGC TTCACCAACC TGTGGCCCT CATCAACGAG	480
55	GCGCTGCTGC ATGATGAGCC CCATGATCAC AAGCTCTCAG ATCAACGGGA GGCCCTGAGT	540
	CATGGCCAGA ACCCTCTGCC CATCTACTGT GCCCTCAACA CCAAAGGGCA GAGCCTGACC	600
	ACTTTTGAAT TTGGGGAGTG GTGCGAGTTC TCTCCCTACG AGGTGGCTT CCCCAAGTAC	660
60	GGGGCCTTCA TCCCCCTCTGA GCTCTTGGC TCCGAGTTCT TTATGGGCA GCTGATGAAG	720

230

	AGGCTTCCTG AGTCCCGCAT CTGCTTCATA GAAGGTATCT GGAGCAACCT GTATGCAGCC	780
5	AACCTCCAGG ACAGCTTATA CTGGGCCTCA GAGCCCAGCC AGTTCTGGGA CCGCTGGTC	840
	AGGAACCAGG CCAACCTGGA CAAGGAGCAG GTCCCCCTTC TGAAGATAGA AGAACCAACCC	900
	TCAACAGCCG GCAGAAATAGC TGAGTTTTC ACCGATCTTC TGACGTGGCG TCCACTGCC	960
10	CAGGCCACAC ATAATTCCT GCGTGGCCTC CATTCCACA AAGACTACTT TCAGCATCCT	1020
	CACTTCTCCA CATGGAAAGC TACCACTCTG GATGGGCTCC CCAACCAGCT GACACCCCTCG	1080
15	GAGCCCCACC TGTGCCTGCT GGATGTTGGC TACCTCATCA ATACCAGCTG CCTGCCCTC	1140
	CTGCAGCCCA CTCGGGACGT GGACCTCATC CTGTCATTGG ACTACAACCT CCACGGAGCC	1200
	TTCCAGCAGT TGCAGCTCCT GGGCCGGTTC TGCCAGGAGC AGGGGATCCC GTTCCCACCC	1260
20	ATCTCGCCCA GCCCCGAAGA GCAGCTCCAG CCTCGGGAGT GCCACACCTT CTCCGACCCC	1320
	ACCTGCCCCG GAGCCCCCTGC GGTGCTGCAC TTTCCTCTGG TCAGCGACTC CTTCCGGGAG	1380
25	TACTCGGCCC CTGGGGTCCG GCGGACACCC GAGGACCGGG CAGCTGGGA CGTGAACCTG	1440
	TCTTCATCGG ACTCTCCCTA CCACTACACG AAGGTGACCT ACAGCCAGGA GGACGTGGAC	1500
	AAGCTGCTGC ACCTGACACA TTACAATGTC TGCAACAACC AGGAGCAGCT GCTGGAGGCT	1560
30	CTGCGCCAGG CAGTGCAGCG GAGGCGGCAG CGCAGGCCCC ACTGATGCC GGGGCCCTG	1620
	CCACCCCTAA CTCTCATTCA TTCCCTGGCT GCTGAGTTC AGGTGGAAC TGTCATCACG	1680
35	CAGTGCTTCA GAGCCTCGGG CTCAGGTGGC ACTGTCCCAG GGTCCAGGCT GAGGGCTGG	1740
	AGCTCCCTTG CGCCTCAGCA GTTTCAGTG GGGTAAGGAG GCCAAGCCCA TTGTGTAAAT	1800
	CACCCAAAAC CCCCGGGCT GTGCCTGTT TCCCTCTGC GCTACCTTGA GTAGTTGGAG	1860
40	CACTTGATAAC ATCACAGACT CATAACAAATG TGAGGCCCTG AGAAAAAAA AAAAAAAA	1920
	CTCGA	1925

45

(2) INFORMATION FOR SEQ ID NO: 87:

50	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1818 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:	
	CGGGGCCCTT CCNCNGNNTT TTTTTTTTTT TTTTTTTTK TATGAGTCTG TRATGTATCA	60
60	AGTGCTCCAA CTACTCAAGG TAGCGCAGAA GGGAAACAG GCACAGGCCG GGGGGTTTG	120

	GGTGATTACA CAAATGGGCT TGGCCTCCCT ACCCCACTGC AAAACTGCTGA GGCGCAAGGG	180
	AGCTCCCAGC CCTCAGCCTG GACCTGGGA CAGTGCCACC TGAGCCCGAG GCTCTGNAAG	240
5	CACTGCGTGA TGACAGITCC CACCTGCAAC TCAGCAGCCA GGGAAATGAAT GAGAGTTAGG	300
	GGTGGCAGGG GCCCCGGCCA TCAGTGGGGC CTGCGCTGCC GCCTCCGCTG CACTGCCTGG	360
10	CGCAGAGCCT CCAGCAGCTG CTCCCTGGTTG TTGCAGACAT TGTAATGTGT CAGGTGCAGC	420
	AGCTTGTCCA CGTCCTCCTG GCTGTAGGTC ACCTTCGTGT AGTGGTAGGG AGAGTCCGAT	480
	GAAGACAGGT TCACCTCCCC AGCTGCCGCC TCCTCGGGTG TCCGCCGGAC CCCAGGGGCC	540
15	GAGTACTCCC GGAAGGGAGTC GCTGACCAGA GGAAAGTGCA GCACCGCAGG GGCTCCGGGG	600
	CAGGTGGGGT CGGAGAAGGT GTGGCACTCC CGAGGCTGGA GCTGCTCTTC GGGGCTGGGC	660
20	GAGATGGGTG GGAACGGGAT CCCCTGCTCC TGGCAGAAC GGGCCAGGAG CTGCAACTGC	720
	TGGAAGGCTC CGTGGAGGTT GTAGTCCAAT GACAGGATGA GGTCCACGTC CCGAGTGGGC	780
	TGCCAGGAGGG GCAGGGAGCT GGTATTGATG AGGTAGCCAA CATCCAGCAG GCACAGGTGG	840
25	GGCTCCGAGG GTGTCAGCTG GTTGGGGAGC CCATCCAGAG TGGTAGCTTT CCATGTGGAG	900
	AAGTGAGGAT GCTGAAAGTA GTCTTGTGG AAATGGAGGC CACGCAGGAA ATTATGTGTG	960
30	GCCTGGCCA GTGGACGCCA CGTCAGAAAGA TCGGTGAAAA ACTCAGCTAT TCTGCCGGCT	1020
	GTTGAGGGTG GTTCTTCTAT CTTCAGAAGG GGGACCTGCT CCTTGTCCAG GTTGGCCTGG	1080
	TTCCCTGACCC AGCGGTCCCA GAACTGGCTG GGCTCTGAGG CCCAGTATAA GCTGTCCCTGG	1140
35	AGGTTGGCTG CATACTGGTT GCTCCAGATA CCTTCTAAGA ACCACATGCG GGACTCAGGA	1200
	AGCCTCTTCA TCAGCTGCC CATAAAAGAAC TCGGAGCCAA AGACCTCAGA GGGGATGAAG	1260
	GCCCCGTACT TGGGGAAGCC GACCTCGTAG GGAGAGAACT CGCACCAACTC CCCAAATTCA	1320
40	AAAGTGGTCA GGCTCTGCC TTTGGTGTTC AGGGCACAGT AGATGGGCAG AGGGTTCTGG	1380
	CCATGACTCA GGGCTCCCG TTGATCTGAG AGCTTGTGAT CATGGGCTC ATCATGCAGC	1440
45	AGCGCCTCGT TGATGAGGGC CCACAGGTTG GTGAAGCAGC TTGGGTAGCC CAAGCGGGCA	1500
	CGCTCGGCCA GCTCCCTGCCG GTACCGCTGC AGCTGGCTGG GGGCCAGCAC ACCCAGCTTG	1560
	TTCTTGGTCA CCTGGGTCTT CAGCAACTCA GTGGGCCCTG CCAGGTCCCTT CTGAGACCAC	1620
50	TCTGGGTCTT YATAAAAGGTT GGCCAAAGGCC CAGGTGGAGC CCGAGGGCCC GGTGATGTAG	1680
	GAGACGCAAT CCAAGAGGCC CCAGCTCCTT TCAGGCCAGC CAGCTGCCCA TACAGGGAAAG	1740
55	TCATTTGCCCG GATCCCACCA CCAGTGGCCA TAATAGCTAC CACTGGGATC TCATCCTCCT	1800
	GCAGGTCTCC ATCCAGCT	1818

(2) INFORMATION FOR SEQ ID NO: 88:

- 5            (i) SEQUENCE CHARACTERISTICS:

  - (A) LENGTH: 539 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

AGGGTAATT	ATATGAAGTG	AAAAAAGTTG	AATGTTCCAG	TCTAAAAGGC	AGTGGGAGAA	60
ATTACATAGC	ATGGAATAA	TAAAATGAAY	TCTTATTAAT	GAGAACGAGG	YTCTTGCAGT	120
GGCAAGTTCT	GCTGGTCACC	CGATGGGAT	GGGAGCCTTT	CAAGCTTTTT	TTTGGGTAAT	180
ACTCACAGTT	TCCAACGTCT	GTGTACTTTT	AAAAATGAGC	TTGTTCTTCC	TTCTGACACT	240
CATCTCAAAG	CTCCATGGTG	ACCGAGAGGT	CTGTTGAAGG	TCACAGGGTC	CTCGCTTGCA	300
TTGGCATACG	GTCCTGTAGC	ATCACTTGTT	AGCCCACITGC	TGCTTGAAGG	AACTAAGAGT	360
ATTCAAGGAT	AGAGAGCTGA	AAATAGGATT	AATMNNITCC	TTTTGACTCT	CCCCTCAAGA	420
TGTCCCTGCT	TTGGTCTGAA	AACCTCTCCT	GACAACCTTT	GCCCCAAAGCA	AACCATCTGC	480
CTTTTCTGAA	CTCTGAGTG	ATATATTAGC	ATCTTCCCTT	CTGAGGCCCTC	GTACTGCCA	539

30

(2) INFORMATION FOR SEQ ID NO: 89:

- 35                   (i) SEQUENCE CHARACTERISTICS:  
                       (A) LENGTH: 855 base pairs  
                       (B) TYPE: nucleic acid  
                       (C) STRANDEDNESS: double  
                       (D) TOPOLOGY: linear

40 SUBJECT DESCRIPTION: SEQ ID NO: 89:

	CCTCTGCCCA	GGCCGCACCC	GAGCTCAGGC	TGTCGCCAC	CCACCAAGTT	CCAGTGC CGC	60
45	ACCAGTGGCT	TATGCGTGCC	CCTCACCTGG	CGCTGCGACA	GGNACTTGGA	CTGCACCGAT	120
	GGCAGCGATG	AGGAGGGAGTG	CAGGATTGAG	CCATGTACCC	AGAAAGGGCA	ATGCCACCG	180
	CCCCCTGGCC	TCCCOCTGCC	CTGCACCGGC	GTCAGTGACT	GCTCTGGGGG	AACTGACAAG	240
50	AAACTGCGCA	ACTGCAGCCG	CCTGGCCTGC	CTAGCAGCGG	AGCTCCGTTG	CACGCTGAGC	300
	GATGACTGCA	TTC CACTCAC	GTGGCGCTGC	GACGGCCACC	CAGACTGTCC	CGACTCCAGC	360
	GACGAGCTCG	GCTGTGGAAC	CAATGAGATC	CTCCCCGAAG	GGGATGCCAC	AACCATGGGG	420
55	CCCCCTGTGA	CCCTGGAGAG	TGTCA CCTCT	CTCAGGAATG	CCACAACCAT	GGGGCCCCCT	480
	GTGAACCTG	GAGAGTGTCC	CCTCTGT CGG	GAATGCCACA	TCC TCCCTTG	CCGGAGACCA	540

GTCTGGAAGC CCAACTGCCT ATGGGGTTAT TGCAGCTGCT GCGGTGCTCA GTGCAAGCCT	600
GGTCACCGCC ACCCTCCCTCC TTTTGTCCCTG GCTCCGAGCC CAGGAGCGCC TCCGCCCACT	660
5 GGGGTTACTG GTGGCCATGA AGGAGTCCCT GCTGCTGTCA GAACAGAAGA CCTCGCTGCC	720
CTGAGGACAA GCACCTGCCA CCACCGTCAC TCAGCCCTGG GCGTACNGSA CAGGAGGAGA	780
10 GCAGTGATGC GGATGGGTAC CGGGCACACC AGCCCTTCAG AGACCTGAGC NCTCTGGCC	840
ACTGGAACCTT CGAAC	855

15

(2) INFORMATION FOR SEQ ID NO: 90:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 628 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

25 AAGGACGTGC CGTGCCGCTG GGTTCTGAGC CGGAGTGGTC GGTGGGTGGG ATGGAGGCGA	60
CCTTGGAGCA GCACCTGGAA GACACAATGA AGAATCCCTC CATTGTTGGA GTCCTGTGCA	120
30 CAGATTCACA AGGACTTAAT CTGGGTTGCC CGGGGACCCCT GTCAGATGAG CATGCTGGAG	180
TGATATCTGT TCTAGCCCAG CAAGCAGCTA AGCTAACCTC TGACCCCCACT GATATTCCCTG	240
35 TGGTGTGTCT AGAACAGAT AATGGGAACA TTATGATCCA GAAACACGAT GGCATCACGG	300
TGGCAGTGCA CAAAATGGCC TCTTGATGCT CATATCTGTT CTTCAGCAGC CTGTCATAGG	360
AACTGGATCC TACCTATGTT AATTACCTTA TAGAACTACT AAAGTTCCAG TAGTTAGGCC	420
40 ATTCAATTAA TGTGCATTAG GCACCTTTCT GTTTATTTAA GAGTCATTG CTTTCTAATG	480
CTCTATGGAC CGACTATCAA GATATTAGTA AGAAAGGATC ATGTTTGAA GCAGCAGGTC	540
45 CAGGTCACTT TGTATATAGA ATTTTGCTGT ATTCAATAAA TCTGTTGGA GGNAAAAAAA	600
AAAAAARAAA AAMTSGAGGG CGGAAGCT	628

50

(2) INFORMATION FOR SEQ ID NO: 91:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1053 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

60

	CTCTTTCTG CAGTTCAAGG GAAAGACGAG ATCITGCACA AGGCACITCG CTTCTGCCCT	60
	TGGCTGGGA AGGGTGGCAT GGARCCCTCTC CGGCTGCTCA TCTTACTCTT TGTCACAGAG	120
5	CTGTCCGGAG CCCACAACAC CACAGTGTTC CAGGGCGTGG CGGGCCAGTC CCTGCAGGTG	180
	TCTTGCCCCAT ATGACTCCAT GAAGCACTGG GGGAGGCGCA AGGCCTGGTG CCGCCAGCTG	240
10	GGAGAGAAGG GCCCATGCCA GCGTGTGGTC ACCACGCACA ACTTGTGGCT GCTGTCCCTTC	300
	CTGAGGAGGT GGAATGGGAG CACAGCCATC ACAGACGATA CCCTGGGTGG CACTCTCAC	360
	ATTACGCTGC GGAATCTACA ACCCCATGAT GCGGGTCTCT ACCAGTGCCA GAGCCTCCAT	420
15	GGCAGTGAGG CTGACACCT CAGGAAGGTG CTGGTGGAGG TGCTGGCAGA CCCCCTGGAT	480
	CACCGGGATG CTGGAGATCT CTGGTTCCTC GGGGAGTCTG AGAGCTTCGA GGATGCCAT	540
	GTGGAGCACA GCATCTCCAG GAGCCTCTTG GAAGGAGAAA TCCCCTTCCC ACCCACTTCC	600
20	ATCCCTCTCC TCCTGGCCTG CATCTTCTC ATCAAGATTC TAGCAGGCCAG CGNCCTCTGG	660
	GCTGCAGCCT GGCATGGACA GAAGCCAGGG ACACATCCAC CCAGTGAACCT GGACTGTGGC	720
25	CATGACCCAG GGTATCAGCT CCAAACCTG CCAGGGCTGA GAGACACGTG AAGGAAGATG	780
	ATGGGAGGAA AAGCCCAGGA GAAGTCCCAC CAGGGACCAG CCCAGCCTGC ATACTTGCCA	840
	CTTGGCCACC AGGACTCCTT GTTCTGCTCT GGCAAGAGAC TACTCTGCCT GAACACTGCT	900
30	TCTCCTGGAC CCTGGAAGCA GGGACTGGTT GAGGGAGTGG GGAGGTGGTA AGAACACCTG	960
	ACAACTCTG AATATTGGAC ATTTTAAACA CTTACAAATA AATCCAAGAC TGTCAATT	1020
35	AAAAAAAAAAAA AAAACGAGGG GGC	1053

## 40 (2) INFORMATION FOR SEQ ID NO: 92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1075 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

50	GCACGAGCCT GATCCTCTCT TTTCTGCAGT TCAAGGGAAA GACGAGATCT TGCACAAGGC	60
	ACTCTGCTTC TGCCCTTGGC TGGGAAGGG TGGCATGGAG CCTCTCCGGC TGCTCATCTT	120
55	ACTCTTGTTC ACAGAGCTGT CCGGAGCCCA CAACACCACA GTGTTCCAGG GCGTGGCGGG	180
	CCAGTCCCTG CAGGTGTCTT GCCCCTATGA CTCCATGAAG CACTGGGGGA GGCGCAAGGC	240
	CTGGTCCCGC CAGCTGGAG AGAAGGGCCC ATGCCAGCGT GTGGTCAGCA CGCACAACTT	300
60	GTGGCTGCTG TCCTTCTGA GGAGGTGGAA TGGGACCACA GCCATCACAG ACGATAACCT	360

	GGGTGGCACT CTCACCATT A CGCTCCGAA TCTACAAACCC CATGATGCCG GTCTCTACCA	420
5	GTGCCAGAGC CTCCATGGCA GTGAGGCTGA CACCCCTCAGG AAGGTCCCTGG TGGAGGTGCT	480
	GGCAGACCCC CTGGATCACCC GGGATGCTGG AGATCTCTGG TTCCCCGGGG AGTCTGAGAG	540
	CTTCGAGGAT GCCCATGTGG AGCACAGCAT CTCCAGGAGC CTCTTGGAAAG GAGAAATCCC	600
10	CTTCCCACCC ACTTCCATCC TTCTCCTCCT GGCTGCATC TPTCTCATCA AGATTCTAGC	660
	AGCCAGCGCC CTCTGGCTG CAGCCTGGCA TGGACAGAAG CCAGGGACAC ATCCACCCAG	720
15	TGAACTGGAC TGTGGCCATG ACCCAGGGTA TCAGCTCAA ACTCTGCCAG GGCTGAGAGA	780
	CACGTGAAGG AAGATGATGG GAGGAAAAGC CCAGGAGAAG TCCCACCAGG GACCAGCCCA	840
	GCCTGCATAC TTGCCACTTG GCCACCAGGA CTCCCTGTTTC TGCTCTGGCA AGAGACTACT	900
20	CTGCCTGAAC ACTGCTTCTC CTGGACCCCTG GAAGCAGGGGA CTGGTTGAGG GAGTGGGGAG	960
	GTGGTAAGAA CACCTGACAA CTTCTGAATA TTGGACATTT TAAACACTTA CAAATAAATC	1020
25	CAAGACTGTC ATATTTAAAAA AAAAAAAAAA AAAAAAAACN CGAGGGGGGN CCCGG	1075

30 (2) INFORMATION FOR SEQ ID NO: 93:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2492 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

40	TCCCGACTCA GCTTCCCACC CTGGGCTTTC CGAGGTGCTK TCGCCGCTGT CCCCACCACT	60
	GCAGCCATGA TCTCCCTAAC GGACACGCAG AAAATTGGAA TGGGATTAAC AGGATTGGAA	120
	GTGTTTTTCC TGTCTTTGG AATGATTCTC TTTTTTGACA AAGCACTACT GGCTATTGGA	180
45	AATGTTTTAT TTGTAGCCGG CTTGGCTTTT GTAATTGGTT TAGAAAGAAC ATTCAAGATTC	240
	TTCTTCCAAA AACATAAAAT GAAAGCTACA GGTTTTTTTC TGGGTGGTGT ATTGTAGTC	300
50	CTTATTGGTT GGCTTTGAT AGGCATGATC TTCGAAATTT ATGGATTTTT TCTCTTGTTC	360
	AGGGGCTTCT TTCTCTCGT TGTGGCTTT ATTAGAAGAG TGCCAGTCCT TGGATCCCTC	420
	CTAAATTTAC CTGGAATTAG ATCATTTGTA GATAAAGTTG GAGAAAGCAA CAATATGGTA	480
55	TAACAAACAAG TGAATTGAA GACTCATTTA AAATATTGAG TTATTATATAA AGTCATTTGA	540
	AGAATATTCA GCACAAAATT AAATTACATG AAATAGCTTG TAATGTTCTT TACAGGAGTT	600
60	AAAAACGTAT AGCCTACAAA GTACCAGCAG CAAATTAGCA AAGAAGCAGT GAAAACAGGC	660

	TTCTACTCAA GTGAACTAAG AAGAACAG CAAGCAAAC GAGAGAGGTG AAATCCATGT	720
	TAATGATGCT TAAGAAAATC TTGAAGGCTA TTTGTGTTGT TTTTCCACAA TGTGCAGAAC	780
5	TCAGCCATCC TTAGAGAACT GTGGTGCGCTG TTTCTTTCT TTTTATTTTG AAGGCTCAGG	840
	AGCATCCATA GGCATTTGCT TTTTAGAAAT GTCCACTGCA ATGGCAAAAA TATTTCCAGT	900
10	TGCACTGTAT CTCTGGAAGT GATGCATGAA TTCGATTGGA TTGTGTCATT TTAAAGTATT	960
	AAAACCAAGG AAACCCAAT TTTGATGTAT GGATTACTTT TTTTGTAAA CATGGTAAA	1020
	ATAAAAACTTC TGTGGTTCTT CTGAATCTTA ATATTCAAA GCCAGGTGAA AATCTGAAC	1080
15	AGATATTCCT TGTTGGAATA TGCAAAGGTG ATTCTTTACT AACTTTTAGT TACTAAATTA	1140
	TAGCTAAGTT TTGTCAGCGAG CATACTCCGG AAAGTCTCAT ACTTCCTGGG AGTCTGCCCT	1200
20	CCTAAGTATC TGTCTATATC ATTCAATTACG TGTAAGTATT TAACAAAAAA GCATTCTTGA	1260
	CCATGAATGA AGTAGTTTGT TTCATAGCTT GTCTCATTGA ATAGTATTAT TGAAGATACT	1320
	AAATGATGCA AACCAAATGG ATTTTTTCCA TGTCACTGATG TAATTTTCTT TTCTCTTTC	1380
25	TTTTTTTAA ATTTTACAG TGGCTTATTA TTTGTTTTC ATAAATTAAA ATAACTTTG	1440
	ATAATGTTTA CTTTAACACA TGTAACATGT TAAAAGGTAA AACTTATGGC TGTTTTAAA	1500
30	GGGCTATTCA TTTAACATGA GTTTCCCTT ATTTTCAGCT TTTTCCTAGC ATATAATAGT	1560
	CATTAAGCAT GACATATCCT TCATATGATC ACTCATCTTG AGTTAATTAG AAAATACCTG	1620
	AGTTCACGTG CTAAAGTCAT TTCACTGTAA TAAACTGACT RTGGTTCTT AAGAACATGA	1680
35	CACTAAAAAA AAAGTGGTTT TTTCCACCG TTGCTGATTA TTAGACAGTA GGAAATAGCT	1740
	GTTTTCTTTA GTTTTACAAG ATGTGACAGC TTTAGTGGTA GATGTAGGGA AACATTTCAA	1800
40	CAGCCATAGT ACTATTGTGTT TTACCACTGA TTGCACTGTT TTGTTTTTT AACAGTTGCA	1860
	AAGCTTTTA ATGCATAAAA GTATAATTGA AATCTGTGGT ATTTATTTAC AAACATGTCT	1920
	ACAAAAATAG ATTACAGCTT ATTTTATTTT TAGTTAAATC TCTTAATACA CAGAGNACT	1980
45	CCCAATCTTG CTCATCTAAA TAAGGAAAGA CTTGGTGTAT AGTGTGATGG TTTAGTCCTTA	2040
	AGGATTAAGA CATTTTGGT ACTTGATTT GACTTACGAT GTATCTGTGA AAATGGGATG	2100
50	ATATTGACAA ATGGAGACTC CTACCTCAAT AGTTAATGGA ATAATAAGAG GCTACTGTG	2160
	TGTCTAATGT TCTTCAAAAAA AGTAATATCC TCACCTGGAG AGTGTCAAAT ACATACTTTG	2220
	AGGATTGACT TTATATAAGG TGCCCTGTAG AAMTCGTTA CACATATTG TGACCCATAT	2280
55	TATTTACAAT GTCTTGATAA TTCTACCTTT TTAGAGCAAG AATAGTATCT GCTAATGTAA	2340
	GGGACATCTG TATTTAACTC CTTTGTAGAC ATGAATTCT ATCAAATGT TCTTTGCACT	2400
60	GTAACAGAGA TTCTTTTTT CAATAATCTT AATTCAAAGC ATTATTTAGGM CTTGAAAGGG	2460

TTTGRTAATC TCCCCGTCCCT TGGTAAAGGT TG

2492

5

(2) INFORMATION FOR SEQ ID NO: 94:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3058 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

15	ACCCCTAAATC AACAGACAAT GGCATTGTCG AAGAGCAACC TGTTAATGAA ATCATGTTAA	60
	AAATCAAGGT TTGGCTTCAG TTTAAATCAC TTGAGGTATG AAGTTTATCC TGTTTTCCAG	120
20	AGATAAACAT AAGTTGATCT TCCAAAATA CCATCATTAG GACCTATCAC ACAATATCAC	180
	TAGTTTTTTTG TGTTTGTGTTG TTTTTGTTT TTTTTCTTGG TAAAGCCATG CACCACAGAC	240
25	TTCTGGGCAG AGCTGAGAGA CAATGGTCCT GACATAATAA GGATCTTGA TTAACCCCCA	300
	TAAGGCATGT GTGTGTATAC AAATATACTT CTCTTTGGCT TTTCGACATA GAACCTCAGC	360
	TGTTAACCAA GGGAAATAC ATCAGATCTG CAACACAGAA ATGCTCTGCC TGAAATTTC	420
30	ACCATGCCTA GGACTCACCC CATTATCCA GGTCTTCTG GATCTGTTA ATCAATAAGC	480
	CCTATAATCA CTTGCTAAC ACTGGGCTTC ATCACCCAGG GATAAAAACA GAGATCATIG	540
	TCTTGGACCT CCTGCATCAG CCTATTCAAA ATTATCTCTC TCTCTAGCTT TCCACAAATC	600
35	CTAAAATTCC TGTCCCAAGC CACCCAAATT CTCAGATCTT TTCTGGAACA AGGCAGAATA	660
	TAAAATAAAT ATACATTTAG TGGCTTGGGC TATGGTCTCC AAAGATCCTT CAAAAATACA	720
40	TCAAGCCAGC TTCATTCACT CACTTTACTT AGAACAGAGA TATAAGGGCC TGGGATGCAT	780
	TTATTTTATC AATACCAATT TTTGTGCCA TGGCAGACAT TGCTAATCAA TCACAGCACT	840
	ATTTCTTATT AAGCCACTG ATTTCTTCAC AATCCTTCTC AAATTACAAT TCCAAAGAGC	900
45	CGCCACTCAA CAGTCAGATG AACCCAACAG TCAGATGAGA GAAATGAACC CTACTTGCTA	960
	TCTCTATCTT AGAAAGCAA AACAAACAGG AGTTTCCAGG GAGAATGGGA AAGCCAGGGG	1020
50	GCATAAAAGG TACAGTCAGG GGAAAATAGA TCTAGGCAGA GTGCCCTAGT CAGGGACCAC	1080
	GGGCGCTGAA TCTGCAGTGC CAACACCAAA CTGACACATC TCCAGGTGTA CCTCCAACCC	1140
	TAGCCTTCTC CCACAGCTGC CTACAAACAGA GTCTCCCAGC CTTCTCAGAG AGCTAAAACC	1200
55	AGAAATTTCAGC AGACTCATGA AAGCAACCCC CCAGCCTCTC CCCAACCTG CGCATTGTC	1260
	TAATTTTTAG AACACTAGGC TTCTTCTTTC ATGTAGTTCC TCATAAGCAG GGGCCAGAAT	1320
60	ATCTCAGCCA CCTGCAGTGA CATTGCTGGA CCCCTGAAAA CCATTCCATA GGAGAATGGG	1380

	TTCCCCAGGC TCACAGTGTA GAGACATTGA GCCCATCACA ACTGTTTGAT CTGCTGGCAG	1440
5	TCTAAAACAG TCCACCCACC CCATGGCACT GCCGCGTGAT TCCCGCGCCA TTCAGAAGTT	1500
	CAAGCCGAGA TGCTGACGTT GCTGAGCAAS AGATGGTGAG CATCAGTGCA AATGCCACAT	1560
	TCAGCACATC AGTCATATGC CCAGTGCAGT TACAAGATGT TGTTTCGGCA AAGCATTTCG	1620
10	ATGGAATAGG GAACTGCAAA TGTATGATGA TTTTGAAAAG GCTCAGCAGG ATTTGTTCTT	1680
	AAACCGACTC AGTGTGTCAAT CCCCGGTTAT TTAGAATTAC AGTTAAGAAG GAGAAACTTC	1740
15	TATAAGACTG TATGAACAAG GTGATATCTT CATAGTGGGC TATTACAGGC AGGAAAATGT	1800
	TTTAACIGGT TTACAAAATC CATCAAAACT TGTGTCAATT CCTGTAAAAG GCAGGAGACA	1860
	TGTGATTATG ATCAGGAAAC TGACACAAAT TATTGTTTC AGCCCCCGTG TTATTGTCCT	1920
20	TTTGAACGTGTTTTTTA TTAAAGCCAA ATTTGTGTTG TATATATTG TATTCCATGT	1980
	GTTAGATGGA AGCAATTTCCT ATCCAGTGTG AATAAAAAGA ACAGTGTAG TAAATTATTA	2040
25	TAAAGCCGAT GATATTTCAT GCCAGGTTAT TCTACCAAGC TGTGCTTGTG GGTTTTCCC	2100
	ATGACTGTAT TGCTTTTATA AATGTACAAA TAGTTACTGA AATGACGAGA CCCTTGTTG	2160
	CACACCAATAAAGAACCT TGATAAGAAC CATATTCTGT TGACAGCCAG CTCACAGTT	2220
30	CTTGCCTGAA GCTTGGTGCA CCCTCCAGTG AGACACAAAGA TCTCTCTTT ACCAAAGTTG	2280
	AGAACAGAGC TGGTGGATTA ATTAATAGTC TTGATATCTT GGCCATGGGT AACCTCATTG	2340
35	TAACATATCAT CAGAATGGGC AGAGATGATC TTGAAGTGTG ACATACACTA AAGTCCAAC	2400
	ACTATGTCAG ATGGGGTAA AATCCATTAA AGAACAGGAA AAAATAATTAA TAAGATGATA	2460
	AGCAAATGTT TCAGCCCAAT GTCAACCCAG TTAAAAAAA AATTAATGCT GTGAAAATG	2520
40	GTGAAATTAG TTGCAAACAT ATATAAGAC ATATGCAGTA AAAAGTCTGT TAATGCACAT	2580
	CCTGTGGAA TGGAGTGTTC TAACCAATTG CCTTTCTTG TTATCTGAGC TCTCCTATAT	2640
45	TATCATACTC AGATAACCAA ATTAAAAGAA TTAGAATATG ATTTTTAATA CACTTAACAT	2700
	TAAACTCTTC TAACTTTCTT CTTTCTGTGA TAATTACAGAA GATAGTTATG GATCTTCAT	2760
	GCCTCTGAGT CATTGTTATA AAAATCAGT TATCACTATA CCATGCTATA GGAGACTGGG	2820
50	CAAAACCTGT ACAATGACAA CCCTGGAAGT TGCTTTTTT AAAAAAATAA TAAATTCTT	2880
	AAATCAACTC TTTTTCTGG TTGCTGTGTT GTTATAAAGT GCAACGKATT CAAGTCCTCA	2940
55	ATATCCTGAT CATAATACCA TGCTATAGGA GACTGGCAA AACCTGTACA ATGACAACCC	3000
	TGGAAGTTGC TTTTTAAAAA AAATAATAAT TTNTTAATCC AAAAAAANAA AAAANIT	3058

## (2) INFORMATION FOR SEQ ID NO: 95:

## (i) SEQUENCE CHARACTERISTICS:

- 5                   (A) LENGTH: 1099 base pairs  
                  (B) TYPE: nucleic acid  
                  (C) STRANDEDNESS: double  
                  (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

10	GGCTTTGTAG CTGCTCCGCA GCCCAGCCCG GGCGCGCTCG CAGAGTCCTA GGCGGTGCGC	60
	GGCNICCTGC CTCCCTCCCTC CTGGCGGGTC GCGGCCCGCG CCTCCGCGGT GCCTGCCTTC	120
15	GCTCTCAGGT TGAGGAGCTC AAGCTTGGGA AAATGGTGTG CATTCCCTGT ATCGTCATTC	180
	CAGTTCTGCT CTGGATCTAC AAAAAATTCC TGGAGCCATA TATATACCCCT CTGGTTTCCC	240
20	CCTTCGTTAG TCGTATATGG CCTAAGAAAG CAATACAAGA ATCCAATGAT ACAAACAAAG	300
	GCAAAGTAAA CTTTAAGGGT GCAGACATGA ATGGATTACC AACAAAAGGA CCAACAGAAA	360
	TCTGTGATAA AAAGAAAGAC TAAAGAAATT TTCCTAAAGG ACCCCATCAT TTAAAAAAATG	420
25	GACCTGATAA TATGAAGCAT CTTCCTTGTA ATTGTCTCTG ACCTTTTTAT CTGAGACCGG	480
	AATTCAAGGAT AGGAGTCTAG ATATTTACCT GATACTAATC AGGAAATATA TGATATCCGT	540
30	ATTTAAAATG TAGITAGITA TATTTAATGA CCTCATTCCCT AAGTTCTTTT TTCTGTTAATG	600
	TAGCTTTCAT TTCTGTTATT GCTGTTTGAA TAATATGATT AAATAGAAGG TTCTGCCAG	660
	TAGACATTAT GTTACTAAAT CAGCACTTTA AAATCTTTGG TTCTCTAATT CATATGAATT	720
35	TGCTGTTTGC TCTAATTCTT TTGGGCTCTT CTAATTGAG TGGAGTACAA TTTTGTGTTG	780
	AAACAGTCCA GTGAAACTGT GCAGGGAAAT GAAGGGTAGAA TTTTGGGAGG TAATAATGAT	840
	GTGAAACATA AAGATTAAAT AATTACTGTC CAACACAGTG GAGCAGCTTG TCCACAAATA	900
40	TAGTAATTAC TATTTATTGC TCTAAGGAAG ATTAAAAAAA GATAGGGAAA AGGGGGAAAC	960
	TTCTTTGAAA AATGAAACAT CTGTTACATT AATGTCTAAT TATAAAATT TTCTGTTTAC	1020
45	TGCATTTCTT CTGTTCCCTAC AAATGTATTA AACATTCACT TTAACTGGTA AAAAAAAAAA	1080
	AAAAAAAAACCC GGGGGGGGG	1099

50

## (2) INFORMATION FOR SEQ ID NO: 96:

## (i) SEQUENCE CHARACTERISTICS:

- 55                   (A) LENGTH: 1580 base pairs  
                  (B) TYPE: nucleic acid  
                  (C) STRANDEDNESS: double  
                  (D) TOPOLOGY: linear

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

	GGCAGAGACT GGAATCTCTC TTICATGAAAA AATGCAGCCC CTTAACCTCA GTTCGACARA	60
5	GTGCAGCTCC TTCTCTCCAC CCACCACAGT GATTCTCCCTT ATCCTGCTGT GCTTTGAGGG	120
	CCTGCTCTTC CTCATTTCA CATCAGTGAT GTTTGGGACC CAGGTGCACT CCATCTGCAC	180
	AGATGAGACG GGAATAGAAC AATTGAAAAA GGAAGAGAGA AGATGGGCTA AAAAAACAAA	240
10	ATGGATGAAC ATGAAAGCCG TTTTGGCCA CCCCTCTCT CTAGGCTGGG CCAGCCCCCT	300
	TGCCACGCCA GACCAAGGGA AGGCAGACCC GTACCAAGTAT GTGGTCTGAA GGACCCCCAC	360
	CGGCATGGCC ACTCAGACAC AAGTCCACAC CACAGCACTA CCGTCCCAC CGTTCTCATG	420
15	AATGTTAAA TCGAAAAGC AAAACAACTA CTCTTAAAC TTTTTTATG TCTCAAGTAA	480
	AATGGCTGAG CATTGCAGAG AAAAAAAAAA GTCCCCACAT TTTATTTTT AAAAACCATC	540
20	CTTTCGATTT CTTTTGGTGA CCGAWGCTGC TCTCTTTCC TTTTAAATC ACTTCTCTGG	600
	CCTCTGGTTT CTCTCTGCTG TCTGCTGGC ATGACTAATG TAGAGGGCGC TGTCTCGCGC	660
	TGTGCCATT CTACTAATG AGTGGACAT GACGCTGTGC TGGATGGAAT AGTCTGGACA	720
25	CCTGGTGGGG GATGCATGGG AAAGCCAGGA GGGCCCTGAC CTCCCACCTGC CCAGGAGGCA	780
	GTCGGGGCT CCCCGATGGG ACATAAAACC TCACCGAAGA TGGATGCTTA CCCCTTGAGG	840
30	CCTGAGAAGG GCAGGATCAG AAGGGACCTT GGACAGCGA CCTCATCCCC CAAGTGGACA	900
	CGGTTGCCT GCTAACCTGC AAAGCAATTG CCTGCCCTGT ACTTTATGGG CTTGGGGTGT	960
	GTAGAATGAT TTTGCGGGGG AGTGGGAGA AAGATGAAAG AGGTCTTATT TGTATTCIGA	1020
35	ATCAGCAATT ATATTCCTG TGATTATTTG GAAGAGTGTG TAGGAAAGAC GTTTTCCAG	1080
	TTCAAAATGC CTTATACAAT CAAGAGGAAA AAAAATTACA CAATTCAGG CAAGCTACGT	1140
40	TTTCCCTTGT TTCATCTGCT TCCCTCTCTCA CCACCCACATC TCCCTCTCTT CCCCAGCAAG	1200
	ATGTCAATTG AGCAGTGTGA ATTCTGACTG CAATAGGCAC CAGTCCCCAA CACATACAGC	1260
	CCCACCATCA TCCCCTCTC ATTTTATAAA CCTCAAAGTG GATTCACCTT CTGATAGTTA	1320
45	ACCCCCATAA ATGTGCACGT ACCITGTGTCT TATCTATATT TTAACCKGGG AGACTGTGT	1380
	CCTGGGCATG GGAGATGACC ATGATGCTGG GGTTACCTCA CAGTCCCCAC CCTTTCAAAG	1440
50	TTNGACATAT GGGCCATCCC ATTGGGCCAG GAATTCCACA GGACACACCT AAGGCTGTGG	1500
	GMAYTGGGGG ACAAAATAGAT TTTCCATTTC GAGGAGGGCA CTTTCCCTGT TGTCAGTTC	1560
	TTGTTTGAA GGGAGGTNGG	1580
55		

(2) INFORMATION FOR SEQ ID NO: 97:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 678 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

10	ATATTTTTTT AGGCTAACGT CCAAGATACA GCATTGAGGA GGCAGCTATG TCTAATGAGG	60
	GCTCTCTTGT TTGCTAGAGA TGAGAGAAAT GTATACTAAT CATTITAATT TGTACTTAAA	120
	ATACATTTTA CTAATCATAT TGATTTAAA TATGACAAAT TCTTCTAGTA GATACTAATC	180
15	TTTCCTGGTT ATCATATTGT CCTAGAGAAG CCTAGGTAAA AATGGGTTCC ACCTAGTCTG	240
	TTTGATATAAC ACCTTCCCCC GTCCCCCTCTC CATCCCTGCC AATTGGGCTC TATGCATATT	300
	GACAAGCAAA TAAGAAAACC TTAGGTTCT TGTATTGAA TTTCCAAAAC AATAAAAGGT	360
20	TTTGACTCAA GATTTGCATT CAAGAAGAGG CAGAAATTCT GTCTTATCTT TTTATCATT	420
	TGTGAACITG TGTTCCTCTG TATGCTTAGA AAATTTACA CACAAGGAAT GTTGAAAAAA	480
25	GTGAGAATTG TAGAGTGCTT GGGGGTTTT TATTTGGTCA GTGCTGATGT GTTARGTGTT	540
	TAGGGAAATA ATGCTTCAGG ACCTTTTGA CAACACAGYT TCATGAATGA CYGGGGGATA	600
	TTWAKGTTGT GCTGAGAAAA GGGAGGGAGT GGGCAGTTGG AATGGGGGAC CCTTACCATT	660
30	GGAAAACATG CATTGN	678

35

## (2) INFORMATION FOR SEQ ID NO: 98:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1253 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

45	ACCTCCCTCC CTCTCAGACT GGTCCGAATC CACGCCTAGC CCAGCCACTG CCACTGGGGC	60
	CATGGCCACC ACCACTGGGG CACTGCCTGC CCAGCCACTT CCCTTGTCTG TTCCCAGCTC	120
50	CCTTGCTCAG GCCCAGACCC AGCTGGGCC CCACCGGNAAGTACCCCCA AGAGGCAAGT	180
	NTTGGCCTGA GACGCTCGTC AGTTCTTAGA TCTTGGGGC CTAAAGAGAC CCCCGTCCCTG	240
	CCTCCCTTCT TTCTCTGTCT CTTCCCTCCT TTTAGTCTTT TTCATCCTCT TCTCTTTCCA	300
55	CCAACCCCTCC TGCATCCTTG CCTTGAGCG TGACCGAGAT AGGTCAATCAG CCCAGGGCTT	360
	CAGTCCTCCT TTATTTATAA TGGGTGGGGG CTACCACCCA CCCTGCTGCA GTCTTGTGAA	420
60	GAGTCTGGGA CCTCCCTCTT CCCCCACTTCT CTCTTCCCTC ATTCCCTTCT CTCTCCTTCT	480

	GGCCTCTCAT TTCCCTACAC TCTGACATGA ATGAATTATT ATTATTTTTC TTTTTCTTTT	540
	TTTTTTACA TTTTGTATAG AAACAAATTTC ATTTAAACAA ACTTATTATT ATTATTTTT	600
5	ACAAAATATA TATATGGAGA TGCTCCCTCC CCGTGTGAAC CCCCCAGTGC CCCCCGTGGC	660
	TGNAGTCTGT GGGCCCATTC GGCCAAGCTG GATTCTGTGT ACCTAGTACA CAGGCATGAC	720
10	TGGGATCCCG TGTACCGAGT ACACGACCCA GGTATGTACC AAGTAGGCAC CCTTGGGCC	780
	ACCCACTGGG GCCAGGGTC GGGGGAGTGT TGGGAGCCTC CTCCCCACCC CACCTCCCTC	840
	ACTTCACTGCA ATTCCAGATT GGACATGTC CATAGCCTTG CTGGGAAGG GCCCACTGCC	900
15	AACTCCCTCT GCCCCAGCCC CACCCCTTGGC CATCTCCCTT TGGGAACTAG GGGGCTGCTG	960
	GTGGGAAATG GGAGCCAGGG CAGATGTATG CATTCCCTTA TGTCCCTGTA AATGTGGGAC	1020
20	TACAAGAAGA GGAGCTGCCT GAGTGGTACT TTCTCTTCCT GTTAATCCTC TGGCCAGCC	1080
	TTATGGCAGA ATAGAGGTAT TTTTGGCTA TTTTTGTAAT ATGGCTTCTG GTCAAAATCC	1140
	CTGTGTAGCT GAATTCCCAA GCCCTGCATT GTACAGCCCC CCACTCCCT CACCACCTAA	1200
25	TAAAGGAATA GTTAACACTC AAAAAAAA AAAAAAAA ACTTGAGGGG GGG	1253

30

(2) INFORMATION FOR SEQ ID NO: 99:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 447 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

40	CAAAGAATGA AATTTACAC TCTCCTCTTC TTGGCAGCTG TAGCAGGGC CCTGGTCTAT	60
	GCTGAAGATG CCTCCTCTGA CTCGACGGGT GCTGATCCTG CCCAGGAAGC TGGGACCTCT	120
45	AAGCCTAATG AAGAGATCTC AGGTCCAGCA GAACCAGCTT CACCCCAAGA GACAACCACA	180
	ACAGCCCAGG AGAYTCGGC GGCAGCAGTT CAGGGGACAG CCAAGGTAC CTCAAGCAGG	240
	CAGGAACTAA ACCCCCTGAA ATCCATAGTG GAGAAAAGTA TCTTACTAAC AGAACAAAGCC	300
50	CTTGCAAAAG CAGGAAAAGG AATGCACGGA GGCAGTGCAG GTGGAAAACA ATTCAATCGAA	360
	AATGGAAGTG AATTGACACA AAAATTACTG AAGAAATTCA GTCTATTAAA ACCATGGCA	420
55	TGAGAAGCTG AAAAGAATKG GATCATT	447

60 (2) INFORMATION FOR SEQ ID NO: 100:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 611 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

10	GGTCTGGGGA GGTGACATGT TGGGCTGTGG GATCCCAGCG CTGGGCTGC TCCTGCTGCT	60
	GCAGGSSWTGG GCAGACGGAA ATGGAATCCA GGGATTCTTC TACCCATGGA GCTGTGAGGG	120
15	TGACATATGG GACCGGGGAGA GCTGTGGGGG CCAGGGGGCC ATTCAATGAG CCCAACYTCT	180
	GCCTGCGTCT CCGGTGCTGC TACCGCAATG GGTCTGCTAC CACCAGCGTC CAGACGAAAA	240
	CGTGCGGGAGG AAGCACATGT GGGCGCTGGT CTGGACGTGC AGCGGCTCC TCCTCCTGAG	300
20	CTGCAGCATC TGCTTGTMT GGTGGGCCAA GCGCCGGGAC GTGCTGCATA TGCCCGTTT	360
	CCTGGCGGGGT CCGTGTGACA TGTCCAAGTC CGTCTCGCTG CTCTCCAAGC ACCGAGGGAC	420
25	CAAGAAGACG CCGTCCACGG GCAGCGTGCC AGTCGCCCTG TCCAAAGAGT CCAGGGATGT	480
	GGAGGGAGGC ACCGAGGGGG AAGGGACGGA GGAGGGTGAG GAGACAGAGG GCGAGGAAGA	540
	GGAGGATTAG GGGAGTCCCC GGGGGACTGG TCAATACAGA TACGGTGGAC GGAAAAAAAAA	600
30	AAAAAAAAAA A	611

## 35 (2) INFORMATION FOR SEQ ID NO: 101:

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 609 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

45	GCATTGGTAA AGCTGGCAGT TGAAACCAAGT TGGACGGCCC AGCTTGCCTC TCTTCTGCCT	60
	GAGTGGGCT CTCAGGTCAC TCGTGCCTG CTGGAGGACA GAGGGGCACC TCAGCCGCC	120
50	CCAAGCCCCAG ACCACAGCAA TAAGGTGGC CTGCAGGAGC CGGGGTGGGG GTGGGGGTGG	180
	GGGGRGCAGG ACCCTRARAT GCCACCAGGA CCTGATGGGC CAGGAAGGGC GTGGACATGG	240
	AGGCTGTGTT TACAGTTTTT TTTTTTTTGT TGTTTTGTGTT TTAAAGAATA CAGAAGGAGC	300
55	CAAGCTTTT TGCACTTTGT ATCCAGCTGC AAGCTCAGGG CAGAGTCAAG GGCTGGGTT	360
	GGAAAAACCT GACTCACAGG AATGCATAAT TGACCCCTTGC AGCTACCCAA TAGCCCTTGG	420
60	AGCTGGCACT GAACCAGGCT GCAAGATTTG ACTGCCTTAA AAACACAAAGG CCCTCTAGGC	480

CTGGCAGGGA TGTCCTGTG CCCAGCACTG GGGGCTCGAA GACTGGTTTC TAGCACTACC	540
GGTCACGGCC ATGTCGTCTT AGAAGGGTCC AGAAGATTAT TTTACGTGTA GTCCATT	600
<b>5 AATGTTCTG</b>	<b>609</b>

## 10 (2) INFORMATION FOR SEQ ID NO: 102:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1770 base pairs
  - (B) TYPE: nucleic acid
  - 15 (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

20 ACGGYCCGGA ATCCCGGGTC GACCCACGCG TCCGGAAAT TGAAACTGAG TGGCCCACGA	60
TGGGAAGAGG GGAAAGCCCA GGGGTACAGG AGGCCTCTGG GTGAAGGCAG AGGCTAACAT	120
25 GGGGTTCCGA GCGACCTTGG CCGTTGGCCT GACCATCTT GTGCTGTCTG TCGTCACTAT	180
CATCATCTGC TTCACCTGCT CCTGCTGCTG CCTTTACAAG ACGTGCCGCC GACCACGTCC	240
GGTTGTCACC ACCACCACAT CCACCACTGT GGTGCATGCC CCTTATCCTC AGCCTCCAAG	300
30 TGTGCCGCC AGCTACCTG GACCAAGCTA CCAGGGCTAC CACACCATGC CGCCTCAGCC	360
AGGGATGCCA GCAGCACCCCT ACCCAATGCA GTACCCACCA CCTTACCCAG CCCAGCCCCAT	420
35 GGGCCCACCG GCCTACCACG AGACCCCTGGC TGGAGGAGCA GCGCGCCCT ACCCGGCCAG	480
CCAGCCTCCT TACAACCCCG SCTACATGGA TGCCCCGAAG SGGNCCTCTG AGCATTCCT	540
GGCCTCTYTG GCTGCCACTT GGTTATGTTG TGTGTGTGCG TGARTGGTGT GCAGGGCGGG	600
40 TTCCCTTACGC CCCATGTGTG CTGCTGTGT CCTGCCTGTA TATGTGGCTT CCTCTGATGC	660
TGACAAGGTG GGGAAACAATC CTTGCCAGAG TGGGCTGGGA CCAGACTTTG TTCTCTTCCT	720
CACCTGAAAT TATGCTTCCT AAAATCTCAA GCCAAACTCA AAGAATGGGG TGGTGGGGGG	780
45 CACCCCTGTGA GGTGGCCCT GAGAGGTGGG GGCCTCTCCA GGGCACATCT GGAGTTCTTC	840
TCCAGCTTAC CCTAGGGTGA CCAAGTAGGG CCTGTCACAC CAGGGTGGCG CAGCTTTCTG	900
50 TGTGATGCAG ATGTGTCTG GTTTCGGCAG CGTAGGCCAGC TGCTGCTTGA GGCCATGGCT	960
CGTCCCCCGGA GTTGGGGTA CCCGTTGCAG AGCCAGGGAC ATGATGCAGG CGAAGCTTGG	1020
55 GATCTGGCCA AGTTGGACTT TGATCCTTGT GGCAGATGTC CCATTGCTCC CTGGAGCCTG	1080
TCATGCCCTGT TGGGATCAG GCAGCCTCCT GATGCCAGAA CACCTCAGGC AGAGCCCTAC	1140
TCAGCTGTAC CTGTCTGCCT GGACTGTCCC CTGTCCCCGC ATCTCCCCTG GGACCAGCTG	1200
60 GAGGGCCACA TGCACACACA GCCTAGCTGC CCCCAGGGAG CTCTGCTGCC CTTGCTGGCC	1260

	CTGCCCTTCC CACAGGTGAG CAGGGCTCCT GTCCACCAGC ACACTCAGTT CTCTTCCCTG	1320
5	CAGTGTTCATTTT AGCCAAACAT TTTGCCTGTT TTCTGTTCA AACATGATAG	1380
	TTGATATGAG ACTGAAACCC CTGGGTTGTG GAGGGAAATT GGCTCAGAGA TGGACAACCT	1440
	GGCAACTGTG AGTCCCCTGCT TCCCACACC AGCCTCATGG AATATGCAAC AACTCCTGTA	1500
10	CCCCAGTCCA CGGTGTCTG GCAGCAGGG AACCTGGGCC AATGGGCCAT CTGGACCAAA	1560
	GGTGGGGTGT GGGGCCCTGG ATGGCAGCTC TGGCCCAGAC ATGAATAACCT CGTGTTCCTC	1620
15	CTCCCTCTAT TACTGTTCA CCAGAGCTGT CTTAGCTCAA ATCTGTTGTG TTTCTGAGTC	1680
	TAGGGTCTGT ACACTTGTT ATAATAAATG CAATCGTTG GAAAAAAA AAAAAAAAC	1740
	TCGTAGGGGG GGCCCGTACCA CAATSGCCTA	1770
20		

## (2) INFORMATION FOR SEQ ID NO: 103:

25	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 1832 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:	
	TGTGGCTGAC GTCATCTGGA GGAGATTGTC TTTCTTTTTC TCCAAAAGGG GAGGAAATG	60
35	AAACTGCAGT GGCCCACGAT GGGAAAGAGGG GAAAGCCCAG GGGTACAGGA GGCCCTCTGGG	120
	TGAAGGCAGA GGCTAACATG GGGTTCGGAG CGACCTTGGC CGTTGGCTGA CCATCTTTGT	180
40	GCTGTCCTGTC GTCACTATCA TCATCTGCTT CACCTGCTCC TGCTGCTGCC TTTACAAGAC	240
	GTGCCGCCGA CCACGTCCGG TTGTCACCAAC CACCACATCC ACCACTGTGG TGCAATGCC	300
	TTATCCTCAG CCTCCAAGTG TGCCGCCAG CTACCCCTGGA CCAAGCTACC AGGGCTACCA	360
45	CACCATGCCG CCTCAGCCAG GGATGCCAGC AGCACCCCTAC CCAATGCAGT ACCCACCACC	420
	TTACCCAGCC CAGCCCATGG GCCCACCGGC CTACCCACGAG ACCCTGGCTG GAGGAGCAGC	480
50	CGCGCCCTAM CCCGSCAGCC AGCCTCCCTTA CAACCCGGCC TACATGGATG CCCGAAGCGG	540
	CCCTCTGAGC ATTCCCCGGC CTCTYTGGCT GCCACCTGGT TATGTTGTGT GTGTGGCTRA	600
	GTGGTGTGCA GGCGGGTTC CTTACGCCCTC ATGTGTGCTG TGTGTGTCCA GGCACGGTTC	660
55	CTTACGCCCTC ATGIGTGCTG TGTGTGTCTT GCCTGTATAT GTGGCTTCTT CTGATGCTGA	720
	CAAGTGGGGA ACAATCCCTG CCAGAGTGGG CTGGGACCAAG ACTTTGTTCT CTTCCCTCACC	780
60	TGAAATTATG CTTCTAAAAA TCTCAAGCCA AACTCAAAGA ATGGGGTGGT GGGGGGCACC	840

	CTGTGAGGTG CCCCTGAGA GGTEGGGGCC TCTCCAGGC ACATCTGGAG TTCTTCTCCA	900
	GCTTACCCCA GGGTGACCAA GTAGGGCTG TCACACCAGG GTGGCGCAST TTCTGTGTGA	960
5	TGCAGATGTG TCCTGGTTTC GGCAGCGTAG CCAGCTGCTG CTTGAGGCCA TGGCTCGTCC	1020
	CCGGAGTTGG GGGTACCCGT TGCAGAGCCA GGGACATGAT GCAGGGGAAG YTTGGGATCT	1080
	GGCCAAGTTG GACTTTGATC CTTTGGGAG ATGTCCCCATT GCTCCCTGGA GCCTGTGATG	1140
10	CCTGTTGGG ATCAGGCAGC CTCCCTGATGC CAGAACACCT CAGGCAGAGC CCTACTCAGC	1200
	TGTACCTGTC TGCCCTGGACT GTCCCCCTGTC CCCGCATCTC CCCTGGGACC AGCTGGAGGG	1260
15	CCACATGCAC ACACACCCCA GCTGCCCTCA GGGAGCTCTG CTGCCCTGTC TGGCCCTGCC	1320
	CTTCCCACAG GTGAGCAGGG CTCCTGTCCA CCAGCACACT CAGTTCTCTT CCCTGCAGTG	1380
	TTTCACTTTT ATTITAGCCA AACATTTTGC CTGTTTCTG TTTCAAAACAT GATACTTGAT	1440
20	ATGAGACTGA AACCCCTGGG TTGTGGAGGG AAATTGGCTC AGAGATGGAC AACCTGGCAA	1500
	CTGTGAGTCC CTGCTTCCCG ACACCAGCCT CATGGAATAT GCAACAACTC CTGTACCCCA	1560
25	GTCCACGGTG TTCTGGCAGC AGGGACACCT GGGCCAATGG GCCATCTGGA CCAAAGGTGG	1620
	GGTGTGGGGC CCTGGATGGC AGCTCTGGCC CAGACATGAA TACCTCGTGT TCCTCCTCCC	1680
	TCTATTACTG TTTCACCAGA GCTGTCTTAG CTCAAATCTG TTGTGTTCT GAGTCTAGGG	1740
30	TCTGTACACT TGTTTATAAT AAATGCAATC GTTTNGAAA AAAANANAA AAAAAAAGG	1800
	GGSGGCGCTC TAAAAGGATN CCCNAAGGG GG	1832

35

(2) INFORMATION FOR SEQ ID NO: 104:

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2237 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

	AGTTCCCGGT ACTTTATTAC CAAGGTTGCC ATCGGAACCA GGAATGACAT TACTCACTAT	60
	CAGAATTGAG AAAATTGGTT TGAAAGATGC TGGCAGTGC ATCGATCCCT ATATTACAGT	120
	TAGTGAAAG GATCTGAATG GCATAGACTT AACTCCCTGTC CAAGATACTC CTGTGGCTTC	180
	AAGAAAAGAA GATACATATG TTCATTAA TGTGGACATT GAGCTCCAGA AGCATGTTGA	240
55	AAAATTAAACC AAAGGTGCAG CTATCTTCTT TGAATTCAA CACTACAAGC CTAAAAAAAG	300
	GTTTACCAAGC ACCAAGTGTT TTGCTTTCAT GGAGATGGAT GAAATTAAAC CTGGGCCAT	360
60	TGTAATAGAA CTATACAAGA AACCCACTGA CTTTAAAAGA AAGAAATTGC AATTATTGAC	420

	CAAGAAACCA CTTTATCTTC ATCTACATCA AACTTTGCAC AAGGAATGAT CCTGACATGA	480
5	TGAACCTGGA ACTTCTGTGA ATTTTACCAAC TCACTAGAAA CCATCATAGC TCTGTGTAGC	540
	ATATTACACCC TTCAACAGGC AGGAAGCAAG CCGTACCCAG ACCAGTAGGC CGGACGGAGT	600
	CAATNGCAA GCTGTACCAC AGAATTCAAGA GTCCAGCACA TCACACTGAC GTATAGGACT	660
10	CCTTGGGATA CAGGTTTATT GTAGATTTTG AAACATGTTT TTACTTTCTT ATTAAATTGTG	720
	CAATTAATAG TCTATTTCTT AATTACCAAC TACTCCTACC CTGCTTCCTG GAACAATACT	780
	GTTGTGGGTA GGATGTGCTC ATCTTCAGAC TTAATACAGC AATAAGAATG TGCTAGAGTT	840
15	TACACATCTG TTCACTTTG CTCCAATATG CTCTTTGAC TTAACGTCAA GCTTTGGGTT	900
	GATGTGGGTA GGGTAGTGTCA AAACGTGTTT GAGAGGAATG GGACCAAGTTC TGCTGCCAA	960
20	GAAGGTCTGT CTGGATGTTT ATAGGCAGCA CCTCTGAAGT GGCTAAATT CACCTGATC	1020
	TGATAGTTT CCTGCTTAGA AAGTGTGCCT TGGCCAGATC AGTATCCCAC ATGGGAGTGT	1080
	TCCCTAGGTT GTAGCTGTGA TTGTTTCCAG ATGACCAGAT TGTTTTCTG AAAATGAGCA	1140
25	TATTTTTAGT CATGTCGATT AGCTGTTCTT CTACATCACA TTGTTACTCT TTCTGATGAT	1200
	GATTCTAGGG TTAACATTGG AACCATCTCA AAATAATTAC AAAGTTTTAG ATGGGTTTAC	1260
30	AATGTCTTCT AAACAATGTA ATCTAAAAAT AATTGAGTCA GATGCTAACG AGATACTGCA	1320
	GGCATAACTG CTGTTTTCT GACAACGTGAT TGTGAAACCT TAAACCTGC ATACCTCTTC	1380
	TTACAGTGAG GAGTATGCAA AATCTGGAAA GATATTCTAT TTTTTTTATA TAGGTAGATA	1440
35	GGATCGCCAT TTATTCCTA TTPAGATATA CTGACATTCA TCCATATGAA AATATGCAGG	1500
	TCATTAGCTT ACTATAATT ACCTTTGACT TAATGGGGCA TAAATAAAAC TTTCATAGTA	1560
40	CACATGAGGT GGATATTTGA TACACAGAAC ATTTGCGGTG GGCTTCTGT GGGTTAGATG	1620
	TAAAGCCCCAC ATATTTTAAT ATTCACTATT TTAAATGAGC AATGCATGAG GGGAAATGCAG	1680
	TGTCAGTACC TGGCTATTT TTAAACTAGT GTAATCACCC TAGTCATACC ATTCAGTATG	1740
45	TTTGCTTTTT AAAATAAGTA ACCACAATTAA AGTTGTTGTA GCCCTTGCAC TTCAAGAGAT	1800
	CTAGTCTTTA CTTTCAGTTG TCTGTTAGGT CCATTCTGTT TACTAGACGG ATGTTAATAA	1860
50	AAACTATGCG AGCCTGAATG AATTCTCAGC CAAATTAGT CTTGCTCTC ATCTTGATTG	1920
	GATTAATTCC AAATTCTAAA ATGATTCACT CCACAATAGC TCTAGGGAT GAAGAATTG	1980
	CCTTACTTTG CCCAGTTCTT AAGACTGTGA GTTGTCAAAT CCCTAGACTG TAAGCTCTTC	2040
55	AAGGAGCAAG AGGCGCATTT TCTCCGTGTC ATGTAATT TCTAAGGTGT TTGGCAGCAC	2100
	TCTGTACCCCT GTGGAGTACT CAGTACCTTT TGTTGATGT TGCTGACAAG ACCTGAAAAAA	2160
60	AAATCCCTTA AAAAAAAAC CCATTAAGT GTAGCAAAAC CGAAAAAAA AAAANAAAAA	2220

ACTCGAGACG GGCCCCG

2237

5

(2) INFORMATION FOR SEQ ID NO: 105:

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1822 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GGTCGACCCA CGCGTCCGGA ATTTTCGTAG CAATAAGTTT GTGCATGTAT AGTAATTTC	60
ATTAGCAAGG TTGTAACCTC TGCCCTTTGG GTCAGTGA TTCTCGTGCC CCAGCCTCCC	120
20 GAGTAGCTGG GACTACAGGC ACGTGCCACC ACCGCCAGCT AATTTTATA TTTTTAGTAG	180
AGACGGGGTT TTGCTGTGTT GGCCAGGCTG GTCTCAAACCT CCTGACCTCA AGTAATCCAC	240
25 CTGGCCTGCT CTTTCAATGT CTTAACATGG CATGTCCTTT AGTTTCATTA TTTTCCTACT	300
CCTTGATATGT CAAGAAATTA CATTTCATG GTCTTATGGA GATGCTGTAA ATTGCTTCAG	360
TGAGTGCTTT TCTAACATGC AGACCATTAA CATTTCCTGT TTGCAGCATG CTGTGTGCAA	420
30 ACACTCAGTA ATTTGGAGTA TTCAATTATT TGTTAGGGCT CTTCTTATTT CCAAATGTGC	480
TGAATTGTCT ATTGATGGGA TTTTCAGATC TTTTCATGAG AACTGGAAAT GTAGCTGGT	540
35 GGCACCTACC TAGGTTGCTA CGTAGTGAGT AGACTTTCTC TTGGGTATAG TAAGCCTCAG	600
ACAGCTTCA CTTTATCTA CTTTACTTGT GGAAATAAAA CAGTCATTTT GTTCTGAAAG	660
AATAAGATAG CTTTCTGTAG AGAAGGAATT CCTACCTCTA AAAGCTGCCT TGAGAACTCA	720
40 GAACTGGCAG TTTTCTGAGG TGATTTTAA ATTTCACTAT TAGGGAGAGT CCAGCATTIG	780
CTGACACAGA TTCTACATAA CTAATGTATG ATAGCAAATG CAAACTATT ATAATGTGGT	840
45 GTATCTTGCCTACACACAGG TTAGAACAAAG TAGACTCTGG CAGCAGATCT CCAGAGACCC	900
AAGTTTAGGT TCTCATAGTG TATTTGAAGT AGTTATACTC CTGGCTTAAG TAGTTTAGTG	960
CCTGGGAGAA TCCATTACTG AAAAGCATTT AACTAAAAA AAAAAAAAAA AAAAAAAAAA	1020
50 AAACCTCGTG CGGAATTCTGG CACGAGCTAA CCCAGAAACA TCCAATTCTC AAACTGAAGC	1080
TCGCACACTCTC GCCTCCAGCA TGAAAGCTC TGCCGCCCTT CTGTGCCCTGC TGCTCATAGC	1140
55 AGCCACCTTC ATTCCCCAAG GGCTCGCTCA GCCAGATGCA ATCAATGCC CAGTCACCTG	1200
CTGYTATAAC TTCACCAATA GGAAAGATCTC AGTGCAGAGG CTCGGAGCT ATAGAAGAAT	1260
CACCAGCAGC AAGTGTCCCA AAGAAGCTGT GATCTCAAG ACCATTGTGG CCAAGGAGAT	1320

60

	CTGTGCTGAC CCCAAGCAGA AGTGGGTTCA GGATTCCATG GACCACCTGG ACAAGCAAAC	1380
	CCAAACTCCG AAGACTTGAA CACTCACTCC ACAACCCAAG AATCTGCAGC TAACTTATTT	1440
5	TCCCCTAGCT TTCCCCAGAC ACCCTGTTTT ATTTTATTAT AATGAATTTT GTTGTGTTGAT	1500
	GTGAAACATT ATGCCTTAAG TAATGTTAAT TCTTATTTAA GTTATTGATG TTTTAAGTTT	1560
10	ATCTTCATG GTACTAGTGT TTTTTAGATA CAGAGACTTG GGGAAATTGC TTTTCCTCTT	1620
	GAACCACAGT TCTACCCCTG GGATGTTTTG AGGGCTTTG CAAGAACAT TAATACAAAG	1680
	AATTTTTTTT AACATTCCAA TGCATTGCTA AAATATTATT GTGGAAATGA ATATTTTGTAA	1740
15	ACTATTACAC CAAATAATA TATTTTTGTA CAAAAAAA AAAAAAAA AAAAAAAA	1800
	AAGGGGCCGC TCGAATTAAG CC	1822

20

## (2) INFORMATION FOR SEQ ID NO: 106:

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 1712 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

	CGTGCCCCAG CCTCCCGAGT AGCTGGRACT ACAGGCACGT SCCACCACGC CCAGCTAATT	60
	TTWATATTTT WAGTAGAGAC GGGGTTTTSC TGTKTTGGCC AGGCTGGTCT CAAACTCCTG	120
35	ACCTCAAGTA ATCCACCTGG CCTGCTCTTT TCATGTCTTA ACATGGCATG TCTTTTAGTTT	180
	TCATTATTTT CCTACTCCCT GTATGTCAAG AAATTACATT TTGCATGTCT TATGGAGATG	240
40	CTGTTAATTG CTTCAGTGAG TGCTTTCTA ATCTGCAGAC CATTACATT TCCTGTTTGC	300
	AGCATGCTGT GTGCAAACAC TCAGTAATTG GGAGTATTCA ATTATTTGTT AGGGCTCTTC	360
	CTATTTCCAA ATGTGCTGAA TTGCTTATTG ATGGGATTTT CAGATTTT CATGAGAACT	420
45	GGAAATGTAG CTGGTGGCA CCTACCTAGG TTGCTACGTA GTGAGTAGAC TTTCTCTTGG	480
	GTATAGTAAG CCTCAGACAG CTTTCACTTT TATCTACTTT ACTTGTGGAA ATAAAACAGT	540
50	CATTTTGTTC TGAAAGAATA AGATAGCTTT CTGTAGAGAA GGAATTCTA CCTCTAAAAG	600
	CTGCCTTGAG AACTCAGAAC TGGCAGTTTT CTGAGGTGAT TTTTAATTT CAGTATTAGG	660
55	GAGAGTCCAG CATTGCTGA CACAGATTCT ACATAACTAA TGTATGTAG CAAATGCAA	720
	ACTATTATAA TGTGGTGTAT CTTGCGCATA CACAGGTAG AACAAAGTAGA CTCTGGCAGC	780
	AGATCTCCAG AGACCCAAGT TTAGGTTCTC ATAGTGTATT TGAAGTAGTT ATACTCCTGG	840
60	CTTAAGTAGT TTAGTGCCTG GGAGAATCCA TTACTGAAAA GCATTTAATC TAAAAAAA	900

250

	AAAAAAAAAA AAAAAAAAAC CTCGTGCCGA ATTGGGCACG AGCAGAAACA TCCAATTCTC	960
5	AAACTGAAGC TCGCACTCTC GCCTCCAGCA TGAAAGTCTC TGCGCCCTT CTGTGCCCTGC	1020
	TGCTCATAGC AGCCACCTTC ATTCCCCAAG GGCTCGCTCA GCCAGATGCA ATCAATGCC	1080
	CAGTCACCTG CTGYTATAAC TTCACCAATA GGAAGATCTC AGTGCAGAGG CTCGGAGCT	1140
10	ATAGAAGAAC CACCAGCAGC AAGTGTCCA AAGAAGCTGT GATCTCAAG ACCATTGTGG	1200
	CCAAGGAGAT CTGTGCTGAC CCCAAGCAGA AGTGGGTICA GGATTCATG GACCACCTGG	1260
15	ACAAGCAAAC CCAAACCTCG AAGACTTGAA CACTCACTCC ACAACCCAAG AATCTGCAGC	1320
	TAACCTATTT TCCCCTAGCT TTCCCCAGAC ACCCTGTTTT ATTTTATTAT AATGAATT	1380
	GTTGTTGAT GTGAAACATT ATGCCCTTAAG TAATGTTAAT TCCTATTAA GTTATTGATG	1440
20	TTTTAAGTTT ATCTTCATG GTACTAGTGT TTTTTAGATA CAGAGACTTG GGGAAATTGC	1500
	TTTCCTCTT GAACCACAGT TCTACCCCTG GGATGTTTG AGGGTCTTG CAAGAACAT	1560
25	TAATACAAAG AATTTTTTTT AACATTCCAA TGCATTGCTA AAATATTATT GTGAAATGA	1620
	ATATTTTGTA ACTATTACAC CAAATAAATA TATTTTGTA CAAAAAAA AAAAAAAA	1680
	AAAAAAAAA AAGSGGCCGC TCGAATTAAG CC	1712

30

## (2) INFORMATION FOR SEQ ID NO: 107:

## 35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1969 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

	CCCTCCCTTC CCCTYGCCAC CTACTGAACC CTCTCCGAG GTGCCCGAGC AGCCGTCTGC	60
45	CCAGCCACTC CCTGGGAGTC CCCCCAGAAG AGCCTATTAC ATCTACTCCG GGGGGAGAA	120
	GATCCCCCTG GTGTTGAGCC GGCCCCCTCTC CTCCAACCTG GCCACTCTTC AGCATCTCTG	180
	TCGGAAGACC GTCAACGGCC ACCTGGACTC CTATGAGAAA GTCACCCAGC TGCGGGGCC	240
50	CATTCGGRAG TTCCCTGGACC AGTACGATGC CCCGTTAA GGGTAAAGG GCGCAAAGGG	300
	CATGGGTGG GAGAGGGAC GCAGGCCCT CTCCCTCCGTG GCACATGGCA CAAGCACAAG	360
55	AAGCCAACCA GGAGAGAGTC CTGTAGCTCT GGGGGAAAG AGGGCGGACA GGCCCTCCC	420
	TCTGCCCTCT CCCTGCAGAA TGTGGCAGGC GGACCTGGAA TGTGTTGGAG GGAAGGGGA	480
60	GTACCAACCTG AGTCTCCAGC TTCTCCGGAG ACCCAGCTGT CCTGGTGGGA CGATAGCAAC	540

	CACAAAGTGGAA TTCTCCTTCATTCCTCAGC TTCCCCCTCTGCCTCCAAACA GGGGACACTT	600
	CGGGAATGCT GAAYTAATGAAACTGCCAG GGAATCTTCAAACTTTCCAA CGGAACCTGT	660
5	TTGCTCTTGTG ATTTGGTTAACCTGAGCT GGTTGTGGAG CCTGGGAAAG GTGGAAGAGA	720
	GAGAGGTCTGAGGGCCCCA GGGSTGCCGGCTGGCGAAGG AAATGGTCAC ACCCCCCGCC	780
10	CACCCAGGCAGGATCCGTG TGACATGCT CCTCTCCCTG GCTCCGGGA GAAGGGCTTG	840
	GGGTGACCTG AAGGGAAACCA TCCCTGGTGCC CCACATCCTC TCCTCCGGGN ACAGTCACCG	900
	AAAACACAGG TTCCAAAGTC TACCTGGTGCCTGAGAGCCC AGGGCCCTTC CTCCGTTTTA	960
15	AGGGGGAAGC AACATTTGGA GGGGACGGAT GGGCTGGTCA GCTGGTCTCC TTTTCCTACT	1020
	CATACTATAC CTTCCGTGAC CTGGGTGGAT GGAGCGGGAG GATGGAGGAG ACGGGACATC	1080
20	TTTCACCTCA GGCTCTGGT AGAGAAGACA GGGGATTCTA CTCTGTGCCT CCTGACTATG	1140
	TCTGGCTAAG AGATTCGCCT TAAATGCTCC CTGTCCCATG GAGAGGGACC CAGCATAGGA	1200
	AAGCCACATA CTCAGCCTGG ATGGGTGGAG AGGCTGAGGG ACTCACTGGA GGGCACCAAG	1260
25	CCAGCCCACA GCCAGGGAAG TGGGGAGGGGG GGGCGGAAAC CCATGCCTCC CAGCTGAGCA	1320
	CTGGGAATGT CAGCCCAGTA AGTATTGGCC AGTCAGGCCTCCTGGTCA GAGCAGAGCC	1380
	ACCAGGTCCC ACTGCCCGA GCCCTGCACA GCCCTCCCTC CTGCTGGGT GGGGGAGGCT	1440
30	GGAGGTCTTGGTGGAGGCTG GACTGCTGCC ACCCCGGGTG CTCCCGCTCT GCCATAGCAC	1500
	TGATCAGTGA CAATTACAG GAATGTAGCA GCGATGGAAT TACCTGGAAC ATTTTTGTT	1560
35	TTTGTGTTTG TTTTGTGTTTG TGTTGGGGGG GGCAACTAAA CAAACACAAA GTATTCTGTG	1620
	TCAGGTATTG GGCTGGACAG GGCAGTTGTG TGTTGGGGTG GTTTTTCTCT CTATTTTTT	1680
	GTTTGTGTTCT TGTTTTTAA TAATGTTTAC AATCTGCCTC AATCACTCTG TCTTTTATAA	1740
40	AGATTCACCC TCCAGTCCTC TCTCCTCCCC CCTACTCAGG CCCCTGAGGC TATTAGGAGA	1800
	TGCTTGAAGA ACTCAACAAA ATCCCAATCC AAGTCAAACT TTGCACATAT TTATATTTAT	1860
45	ATTCAGAAAA GAAACATTTC AGTAATTAT AATAAAGAGC ACTATTTTT AATGAAAAAA	1920
	AAAAAAAAAAAAA CGACGCTGGT GACCGGAATY CGACGTACG	1969

50

(2) INFORMATION FOR SEQ ID NO: 108:

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1734 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

	CGGGTCCCAA GCCTGTGCCT GAGCCTGAGC CTGAGCCTGA GCCCAGCCG GGAGCCGTC	60
5	GCGGGGGCTC CGGGCTGTGG GACCGCTGGG CCCCCAGCGA TGGCGACCCCT GTGGGGAGGC	120
	CTTCTTCGGC TTGGCTCCTT GCTCAGCCTG TCGTGCCTGG CGCTTTCCGT GCTGCTGCTG	180
10	GCGCATGTNC AGACGCCGCC AAGAATTTCG AGGATGTCAAG ATGTAATGT ATCTGCCCTC	240
	CCTATAAAGA AAATTCTGGG CATAATTATA ATAAGAACAT ATCTCAGAAA GATTGTGATT	300
15	GCCTTCATGT TGTGGAGCCC ATGCCTGTGC GGGGGCCTGA TGTAGAAGCA TACTGTCTAC	360
	GCTGTGAATG CAAATATGAA GAAAGAAGCT CTGTCACAAT CAAGGTTACC ATTATAATTT	420
20	ATCTCTCCAT TTTGGGCCTT CTACITCTGT ACATGGTATA TCTTACTCTG GTTGAGCCCA	480
	TACTGAAGAG GCGCCTCTTT GGACATGCAC AGTTGATACA GAGTGTGAT GATATTGGGG	540
25	ATCACCAAGCC TTTTGCAAAT GCACACGATG TGCTAGCCCG CTCCCGCAGT CGAGCCAACG	600
	TGCTGAACAA GGTAGAATAT GCACAGCAGC GCTGGAAGCT TCAAGTCCAA GAGCAGCGAA	660
	AGTCTGTCTT TGACCGGCAT GTTGTCCCTCA GCTAATTGGG GAATTGAATT CAAGGTGACT	720
30	AGAAAGAAC AGGCAGACAA CTGGGAAAGA ACTGACTGGG NTTTGCTGG GTTTCATTTT	780
	AATACCTTGT TGATTTCACC AACTGTGCT GGAAGATTCA AAACCTGGAAG CAAAAACTTG	840
35	CTTGATTTT TTTCTTGTG AACGTAATAA TAGAGACATT TTTAAAAGCA CACAGCTCAA	900
	AGTCAGCCAA TAAGTCTTT CCTATTGTG ACTTTTACTA ATAAAATAA ATCTGCCTGT	960
	AAATTATCTT GAAGTCTTTT ACCTGGAACA AGCACTCTCT TTTTCACCCAC ATAGTTTAA	1020
40	CTTGACTTT TCAGGGTTTT TGTGTGTTGTT GTTTTTGTT TGTTTGTGTT	1080
	GGTGGGAGAG GGGAGGGATG CCTGGGAAGT GGTTAACAAAC TTTTTCAAG TCACTTTACT	1140
45	AAACAAACTT TTGTAATAG ACCTTACCTT CTATTTCGA GTTTCATTAA TATTTTGCAG	1200
	TGTAGCCAGC CTCATCAAAG AGCTGACTTA CTCATTTGAC TTTTGACTG ACTGTATTAT	1260
	CTGGGTATCT GCTGTGTCTG CACTTCATGG TAAACGGGAT CTAAAATGCC TGGTGGCTTT	1320
50	TCACAAAAAG CAGATTTCTC TCATGTACTG TGATGTCTGA TGCAATGCAT CCTAGAACAA	1380
	ACTGCCATT TGCTAGTTA CTCTAAAGAC TAAACATAGT CTGGTGTGTT GTGGTCTTAC	1440
55	TCATCTCTA GTACCTTAA GGACAAATCC TAAGGACTTG GACACTTGCA ATAAAGAAAT	1500
	TITATTTAA ACCCAAGCCT CCCTGGATTG ATAATATATA CACATTGTCA AGCATTCCG	1560
	GTCGTGGTGA GAGGCAGCTG TTTGAGCTCC AATGTGTGCA GCTTTGAACG AGGGCTGGGG	1620
	TTGTGGGTGC CTCTCTGAA AGGTCTAACCTT ATTATTGGAT AACTGGCTTT TTTCTTCCTC	1680
	TTTGGAAATGT AACAAATAAAA ATAATTTTG AAACATCAAA AAAAAAAA AAAA	1734

## (2) INFORMATION FOR SEQ ID NO: 109:

## 5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2003 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

CGCAGGGGGC	GCGCGGCCCG	GGGACTCGCA	TTCCCCGGTT	CCCCCTCCAC	CCCACGCGGC	60
15 CTGGACCATG	GACGCCAGAT	GGTGGGCAGT	GGTGGTGTG	GCTGCGTTCC	CCTCCCTAGG	120
GGCAGGTGGG	GAGACTCCCG	AAGCCCTCC	GGAGTCATGG	ACCCAGCTAT	GGTTCTTCCG	180
20 ATTGTGGTG	AATGCTGCTG	GCTATGCCAG	NTTTATGGTA	CCTGGCTACC	TCCTGGTGCA	240
GTACTTCAGG	CGGAAGAACT	ACCTGGAGAC	CGGTAGGGC	CTCTGCTTTC	CCCTGGTGAA	300
AGCTTGTGTG	TTTGGCAATG	AGCCCAAGGC	CTCTGATGAG	GTTCCCTGG	CGCCCCGAAC	360
25 AGAGGGGGCA	GAGACCACCC	CGATGTGGCA	GGCCCTGAAG	CTGCTCTTCT	GTGCCACAGG	420
GCTCCAGGTG	TCTTATCTGA	CTTGGGGTGT	GCTGCAGGAA	AGAGTGTGATGA	CCCGCAGCTA	480
30 TGGGGCCACA	GCCACATCAC	CGGGTGAGCG	CTTTACGGAC	TCGCAGTTCC	TGGTGCTAAT	540
GAACCGAGTG	CTGGCACTGA	TTGTGGCTGG	CCTCTCTGT	GTTCTCTGCA	AGCAGCCCCG	600
GCATGGGGCA	CCCATGTACC	GGTACTCCTT	TGCCAGCCTG	TCCAATGTGC	TTAGCAGCTG	660
35 GTGCCAATAC	GAAGCTCTTA	AGTTCGTCA	CTTCCCCACC	CAGGTGCTGG	CCAAGGCCCTC	720
TAAGGTGATC	CCTGTCATGC	TGATGGAAA	GCTTGTGTCT	CGGGCGANTA	ACGAACACTG	780
40 GGAGTACCTG	ACAGCCACCC	TCATCTCCAT	TGGGGTCAGC	ATGTTTCTGC	TATCCAGCGG	840
ACCAAGAGCCC	CCGAGCTCCC	CAGCCACAC	ACTCTCAGGC	CTCATCTTAC	TGGCAGGTTA	900
TATTGCTMTT	GACAGCTTCA	CCTCAAACIG	GCAGGATGCC	TGTTTGCTA	TAAGATGTCA	960
45 TCGGTGCAGA	TGATGTTTGG	GGTCAATTTC	TTCTCTGCC	TCTTCACAGT	GGGSTCACTG	1020
CTAGNAACAG	GGGGGMCTA	CTGGAGGGAA	CCCGCTTCAT	GGGGCGACAC	AGTGAGTTG	1080
50 CTGCCCATGC	CCTGCTACTC	TCCATCTGCT	CCGCATGTGG	CCAGCTCTTC	ATCTTTTACA	1140
CCATTGGGCA	GTITGGGGCT	GCCGTCTTCA	CCATCATCAT	GACCCCTCCGC	CAGGCCTTGT	1200
CCATCCTTCT	TTCTCTGCCTT	CTCTATGGCC	ACACTGTCAC	TGTGGTGGGA	GGGCTGGGG	1260
55 TGGCTGTGGT	CTTGTGCTGCC	CTCCTGCTCA	GAGTCTACGC	GGGGGGCCGT	CTAAAGCAAC	1320
GGGGAAAGAA	GGCTGTGCCT	GTGAGTCTC	CTGTGCAGAA	GGTTTGAGGG	TGGAAAGGGC	1380
60 CTGAGGGGTG	AAATGAAATA	GGACCCCTCCC	ACCATCCCT	TCTGCTGTAA	CCTCTGAGGG	1440

	AGCTGGCTGA AAGGGCAAAA TGCAGGTGTT TTCTCAGTAT CACAGACCAG CTCTGCAGCA	1500
	GGGGATTGGG GAGCCCAGGA GGAGCCTTC CCTTTGCCT TAAGTCACCC ATCTTCAGT	1560
5	AAGCAGTTA TTCTGAGCCC CGGGGGTAGA CAGTCCTCAG TGAGGGTTT TGGGGAGTT	1620
	GGGGTCAAGA GAGCATAGGT AGGTTCCACA GTTACTCTTC CCACAAGTTC CCTTAAGTCT	1680
10	TGCCCTAGCT GTGCTCTGCC ACCCTCCAGA CTCACTCCCC TCTGCAAATA CCTGCATTTC	1740
	TTACCCCTGGT GAGAAAAGCA CAACCGGTGT AGGCTCCAAT GCTGCTTTC CAGGAGGGTG	1800
	AAGATGGTGC TGTGCTGAGG AAAGGGGATG CAGAGCCCTG CCCAGCACCA CCACCTCCTA	1860
15	TGCTCCCTGGA TCCCTAGGCT CTGTTCCATG AGCCTGTTGC AGGTTTTGGT ACTTTAGAAA	1920
	TGTAACCTTT TGCTCTTATA ATTTTATTT ATTAAATTAA ATTACTGCAA AAAAAAAA	1980
20	AAAAAAAAATCG GGGGGGGGCC CGN	2003

## (2) INFORMATION FOR SEQ ID NO: 110:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1320 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 30 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

35	GCTGAGCTGC CTTGAGGTGC AGTGTGTTGGG ATCCAGAGCC ATGTCGGACC TGCTACTACT	60
	GGGCTGATT GGGGGCCTGA CTCTCTTACT GCTGCTGAGG CTGCTGGCT TTGCCCCGGTA	120
	CTCAGGGCTA CTGGCTGGGG TGGAAAGTGAG TGCTGGGTCA CCCCCCATCC GCAACGTCAC	180
40	TGTGGCCTAC AAGTCCACA TGGGCTCTA TGGTGAGACT GGGGGCTTT TCACTGAGAG	240
	CTGCAGCATC TCTCCAAGC TCCGCTCCAT CGCTGTCTAC TATGACAACC CCCACATGGT	300
	GCCCCCTGAT AAGTGGCGAT GTGCCGTGGG CAGCATCCTG AGTGAACGTG AGGAATCGCC	360
45	CTCCCCCTGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CCTTCCCGGC	420
	ACCCAGCCAT GTGGTGACAG CCACCTTCCC CTACACCACC ATTCTGTCCA TCTGGCTGGC	480
50	TACCCGCCGT GTCCATCCTG CCTTGGACAC CTACATCAAG GAGCGGAAGC TGTGTGCCATA	540
	TCCTCGGCTG GAGATCTACC AGGAAGACCA GATCCATTTC ATGTGCCAC TGGCAGGGCA	600
	GGGAGACTTC TATGTGCCCTG AGATGAAGGA GACAGAGTGG AAATGGCGGG GGCTTGTGGA	660
55	GGCCATTGAC ACCCAGGTGG ATGGCACAGG AGCTGACACA ATGAGTGACA CGAGTTCTGT	720
	AAGCTTGGAA GTGAGCCCTG GCAGCCGGGA GACTTCAGCT GCCACACTGT CACCTGGGC	780
60	GAGCAGCCGT GGCTGGGATG ACGGTGACAC CCGCAGCGAG CACAGCTACA GCGAGTCAGG	840

	TGCCAGCGC TCCCTTTTG AGGAGCTGGA YTTGGAGGGC GAGGGGCCT TAGGGGAGTC	900
5	ACGGCTGGAC CCTGGACTK AGCCCCCTGGG GACTACCAAG TGGCTCTGGG AGCCCACITGC	960
	CCCTGAGAAG GGCAAGGAGT AACCCATGGC CTGCACCTCTC CCTGCAGTGC AGTTGCTGAG	1020
	GAACTGAGCA GACTCTCCAG CAGACTCTCC AGCCCTCTTC CTCCTTCCTC TGGGGGAGGA	1080
10	GGGGTTCTG AGGGACCTGA CTCCCCCTGC TCCAGGCCTC TTGCTAAGCC TTCTCCTCAC	1140
	TGCCCTTAG GCTCCCAGGG CCAGAGGAGC CAGGGACTAT TTTCTGCAAC CAGCCCCCAG	1200
15	GGCTGCCNCC CCTGTGTGTGT CTTTTTTCA GACTCACAGT GGAGCTTCCA GGACCCAGAA	1260
	TAAAGCCAAT GATTACTTG TTTCAAAAAA AAAAWAAAAA AAAAAAAA AAAAAAAA	1320

20 (2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1962 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

30	CGGACCCCTT CCTCCCTCCTC NAAGCATGTC CCACCATTGT GGCAGGGCT GGGGANACAG	60
	TCACCTGATG CGGGGACCAAC GGCCACTCCA CCTCGSTGGC GCTGTCAGTG GGCAGCACTG	120
35	GCTGGGCCTG CACTGAGGTC CCTGCTGGGG CAGTTCTTCC AGAATTATCT TCAGAGGGGG	180
	CCTCCAGCTC CCTGGTACCC TCAGGGGCC CTTGGCTGG AACCAGGGAA GGGGCACCCCT	240
40	CGGAGCTTCC TGTCCTCTCG CTCTCTCCTC GAGGGACCCC AGATAGCTCA GGACCACCAAG	300
	TTGCCTCCCC CACCTCTCTT GCCTCAACCA GAGTGGAAAGG TGATGGGGAT GCTAGGTTCC	360
	TCTCCCTGGG AGTGGGCAGA GTCTCAGTAG GTGGTCCATG GACCTTGGA GGCCTGGAAAG	420
45	CTTCTGACTC TCCATCAGGA AGTGGTGATG CACCAGGCTG CAGGACTGCC CTTGCTGGCG	480
	CCTGGGAGAG TGACTCCCTCC TGGGCTGCTG GCTCAGTGG GAGAGAGGCC TCAGGGCCCG	540
50	GGCTGCTGAG CTCGCTGGGC CATGCCACACA GAGCCTCATC CTCCACCTCC TCCTCTTCTT	600
	CTTCCTCCCTC TTTCCTTCTC TCATCTTCAT ATTTCCTCTTC TTCTCTCAAT GCCTTACCTT	660
	CCTCTTYTGR AAACCCCGTG GGCGGTACCA TGGATTGTGT TTCAAATTCT AGGAGCGTCC	720
55	TAGGGGCCCTC TGCTGGGTCT TCTGGAGTGG AGCTTCCACC TCCTCCGTCC TCCATGATGG	780
	GGATGGAGTA RATGGCCCCA CGGGATTCAAC TCTCTGTGGC TTCTGTAGGC AGCTGCAGTT	840
60	CCTCCAGGGT CTCTGTCACT GTGACRATAG CCTCTAGTCC ATCAAAAGCT GGGTTGGAGG	900

	CTGGGTTGGA GGCCTCAGGG ATGGCAGAAG GCTGGGCCGA GTCTCGGAAG CAGTARACGT	960
	TGAAGCGGCT GTGCTTATTG GGGAAGCCAG TCTGGTTGGG GAAGANGAAG AGAGTCCTGA	1020
5	CACCAGGCAA GCCCCCACCA CAGCGCTGGC TGGGTGTGAC GATGGGGTAG CGCACANTGC	1080
	CATCAGCTAG CCACCTGGGC TGCAGTGGTC CAGGCCACCA TCCCAGGCTG CATAACAGTG	1140
10	GCCCCGTGGTG GCAATCTCTG CACCCCGCTC CTGGCAGTAC GCCCCGTGCTT CCTCCAATGT	1200
	CAGCTTCTCT GGAGGGTCAC CCAGGAACAG TTCTCCATTT AGGTCTTCAG CATAACAGTA	1260
	CACATCATAG AGGTCATCCG GGTCCACCAC ACCATAGTTC CGGACCCCCGG GGAAGCCATC	1320
15	CATGTCTCCG TAACAGGCCT CTGGTGGGT CTGGATGGGA TACCTTTGAC CTTGAMCTCC	1380
	ACAGCGTCGC TGCTGTCATC GATGCCGTGC TGGACCTCAC AGCGATAGAT ACCTGAGTCG	1440
	TTGGGGCGCA GCTCGCTCAG CGCCAGGGGA GACGTCGGTG AGCGACGCTG GGTACGCAGG	1500
20	CAGTGCCACG CGGAACCGGT AGGCCTCGTT CACCTTGACG CGCACTCCCC GCGCCACCAG	1560
	CACYTCTGCC TCCCCGGCCCC GGGACAGGAA AGTCCACTTG ACCCGCGGAG AGCCCAGCAC	1620
25	AGCCCCGGCGG CTCGGCGGTG SCCGCAGGTA GTGGACGTGG CAAGGGATGK TGAGGGSCC	1680
	GCGGAGCAAC GCCYTGCACT GGCGCGTCGC CGCGCATGCG CACCGCGAAAA GCGCGKTCCCT	1740
30	CTGAGCTGTC TCCCTCCAGA ACATCTGCTA AAGCTGCAGG AGCCTGGGCC AGGACCACGG	1800
	CTGCCAGCAG GGGCAGGAAC AGCTGGGCCA TGCTGCAGGC TACCCAGGGC TGGGGTTGGG	1860
	TCGCGGCAGT GCGAAGTTTG TCGCCTCCTC CGGGGGTCTC CTCCGGTKC ACGGCTCAGT	1920
35	NCCTGCAGCT GCAGCTGAGA CTGGGGCGGA GACTGGGGGA GC	1962

## 40 (2) INFORMATION FOR SEQ ID NO: 112:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1785 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

50	AAGTTTCAGC CAAACTTCGG GCGGCTGAGG CGGCGGCCGA GGAGCGGCCGG ACTCSGGCG	60
	CGGGGAGTCG AGGCATTGTC GCCTGGGCTT CGGAGCGTAC CGCAGGGCCT GAGCCTTGA	120
	AGCAGGAGGA GGGGAGGAGA GAGTGGGGCT CCTCTATCGG GACCCCTCC CCATGTGGAT	180
55	CTGCCAGGC GGCGCGGCCG GCGGAGGAGG CGACCGAGAA GATRCCCGCC CTGGCGCCCC	240
	CTCTGCTGTG GGCGCTGCTG GCGCTCTGGC TGTGCTGCC GACCCCGCGC ATGCATTGCA	300

	GTGTCGAGAT GGCTATGAAC CCTGTGTAAG TGAAGGAATG TGTGTTACCT ACCACAATGG	360
	CACAGGATAC TGCAAAATGTC CAGAAGGCCTT CTGGGGGAA TATIGTCAAC ATCGAGACCC	420
5	CTGTGAGAAG AACCGCTGCC AGAATGGTGG GACTTGTGTG GCCCAGGCCA TGCTGGGAA	480
	AGCCACGTGC CGATGTGCCT CAGGGTTAAC AGGAGAGGAC TGCCAGTACT CGACATCTCA	540
10	TCCATGCTTT GTGTCTCGAC CTTGCCTGAA TGGCGGCACA TGCCATATGC TCAGCCGGAA	600
	TACCTATGAG TGCACCTGTC AAGTCGGTT TACAGGTAAG GAGTCCAAT GGACCGATGC	660
	CTGCCTGTCT CATCCCTGTG CAAATGGAAG TACCTGTACC ACTGTGGCCA ACCAGTTCTC	720
15	CTGCAAATGC CTCACAGGCT TCACAGGGCA GAAGTGTGAG ACTGATGTCA ATGAGTGTGA	780
	CATTCCAGGA CACTGCCAGC ATGGTGGCAC CTGCCTCAAC CTGCCTGGTT CCTACCAGTG	840
20	CCAGTGCCTT CAGGGCTTCAGTA CTGTGACAGC CTGTATGTGC CCTGTGCACC	900
	CTCGCCCTGT GTCAATGGAG GCANCTGTCG GCAGACTGGT GACTTCACIT TTGAGTGCAA	960
	CTGCCCTTCCA GAAACAGTGA GAAGAGGAAC AGAGCTCTGG GAAAGAGACA GGGAAAGTCTG	1020
25	GAATGGAAAAA GAACACGGATG AGAATTAGAC ACTGGAAAAT ATGTATGTGT GGTAAATAAA	1080
	GTGCTTTAAA CTGAATTGAC ATTAACAGTR GGTGATCAAC TTTMCTATGT GCTTGTGCTT	1140
	TTGCTTTGA TGGAGTAATT CATTGTTTC TTATCCACCT AAATGCACCC AGCTGCCCTT	1200
30	GATTTTCCTCT GGGCTACTGG CCTTCACAAAC CCTCTCCCCAT GTACCCCTCTC TGACTTTGGG	1260
	GTAACCCTCC CCTAACCTAA AGCTAGAGAA TTCTGAAACT GAGGAGGGGA TCCTCTGTTA	1320
35	ATCAGTGAGC ACTMMMTGAT GACCTGATAG ATGATATATG AGAGACTATG CGTGGCACAA	1380
	TACTTTGTTA CACTCTTCAC TGATACAAGT GTTCTAGAGT GYACACACAA CCCAAAGATA	1440
40	GAAATAAAAAA GAGGAGGAGT GTCCGGGAGC TTGGGGCTG GTGTTCCATG GAGAGGGAGA	1500
	AAGGAACAAG CTTGRCCAAT TCATTCAACT CCTTATAAAA ATGATGAGGA GGCTGAAAAC	1560
	CAAGAATTIT GATTGGGAAC AGAATACAAG CAGCTGAAKC AGATGAWTTA CTAAGCAACA	1620
45	AAGATCCTGT TTTTATACAA ATATCCTTAG TACAAAAACA AAARAAGGAA AACTGTAGGG	1680
	GGGAGTAATG TGCTAAGTAA GCAGAATTGC CTCCAAAAGA AGTTGTTCT AGTTACTCTT	1740
	TTCCGGGTNG GGATCTTTAG NTTCGGTAT TGTGGGTATG GTTCC	1785
50		

## (2) INFORMATION FOR SEQ ID NO: 113:

55

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1842 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

	GGAGCCTCTC TTGCAACTTC TGCCACCGCG GGCCACCGCG GCGGCCTGAT CCCGCAGAGG	60
5	AAGTCGGCGC CGTGGAGCGA TGACCCCGGG CGGTCCGGC GGGGGCCCGG GGCTGCCACA	120
	GCGCCGCGC CTCTGCTGTC TGCTGCTGCT GCMGCTGTTG TTAGTCACCG CGGAGCCGCC	180
10	GAAACCTGCA GGAGTCTACT ATGCAACTGC ATACTGGATG CCTGCTGAAA AGACAGTACA	240
	AGTCAAAAT GTAATGGACA AGAATGGGA CGCCTATGGC TTTTACAATA ACTCTGTGAA	300
	AACCACAGGC TGGGGCATCC TGGAGATCAG AGCTGGCTAT GGCTCTCAA CCCTGAGCAA	360
15	TGAGATCATC ATGTTTGTGG CTGGCTTTTG GGAGGGTTAC CTCACTGCC CACACATGAA	420
	TGACCACTAC ACAAACCTCT ACCCACAGCT GATCACGAAA CCTTCCATCA TGGATAAAGT	480
20	GCAGGATTTT ATGGAGAAGC AAGATAAGTG GACCCGGAAA AATATCAAAG AATACAAGAC	540
	TGATTCACTT TGGAGACATA CAGGCTATGT GATGGCACAA ATAGATGCC TCTATGAGG	600
	ACCAAAGAAG AGGGCTATAT TAGAAGGGAC AAAGCCAATG ACCCTGTCC AGATTCAAGT	660
25	CCTGAATAGT GTGGAGATC TATTGGATCT GATTCCCTCA CTCTCTCCCA CAAAAAACGG	720
	CAGCCTAAAG GTTTTAAGA GATGGGACAT GGGACATTGC TCCGCTCTTA TCAAGGTTCT	780
30	TCCTGGATTT GAGAACATCC TTTTGCTCA CTCAAGCTGG TACACGTATG CAGCCATGCT	840
	CAGGATATAT AAACACTGGG ACTTCACACRT CATAGATAAA GATACCAAGCA GTAGTCGCCT	900
	CTCTTTCAGC AGTTACCCAG GGTTTTGGA GTCTCTGGAT GATTTTTACA TTCTTAGCAG	960
35	TGGATTGATA TTGCTGCAGA CCACAAACAG TGTGTTTAAT AAAACCTGC TAAAGCAGTA	1020
	ATACCCGAGA CTCTCCTGTC CTGGCAAAGA GTCCGTGTGG CCAATATGAT GGCAGATAGT	1080
40	GGCAAGAGGT GGGCAGACAT CTTTCAAAAA TACAACCTCTG GCACCTATAA CAATCAATAC	1140
	ATGGTTCTGG ACCTGAAGAA AGTAAAGCTG AACCACAGTC TTGACAAAGG CACTCTGTAC	1200
	ATTGTGGAGC AAATTCTTAC ATATGTAGAA TATTCTGAAC AACTGATGT TCTACGGAAA	1260
45	GGATATTGGC CCTCCTACAA TGTTCCCTTC CATGAAAAAA TCTACAACCTG GAGTGGCTAT	1320
	CCACTGTTAG TPCAGAAGCT GGGCTTGGAC TACTCTTATG ATTTAGCTCC ACGAGCCAAA	1380
50	ATTTCCGGC GTGACCAAGG GAAAGTGACT GATACGGCAT CCATGAAATA TATCATGCGA	1440
	TACAACAATT ATAAGAAGGA TCCCTACAGT AGAGGTGACC CCTGTAATAC CATCTGCTGC	1500
	CGTGAGGACC TGAACTCACC TAACCCAAGT CCTGGAGGTT GTTATGACAC AAAGGTGGCA	1560
55	GATATCTACC TAGCATCTCA GTACACATCC TATGCCATAA GTGGTCCCAC AGTACAAGGT	1620
	GGCCTCCCTG TTTTCGCTG GGACCGTTTC AACAAAACTC TACATCAGGG CATGSCAGAG	1680
60	GTCTACAAC TTAGTTTAT TACCATGAAA CCAATTGATG AACTTGATAT AAAATGAAGG	1740

AGGGAGATGA CGGACTAGAA GACTGTAAAT AAGATAACCA AGGCACATT TTAGCTATGT 1800  
 TTTTCCCATC AGAATTATGC AATAAAATAT ATTAATTGT CA 1842

5

## (2) INFORMATION FOR SEQ ID NO: 114:

10

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1960 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

GAATTGGCA CGAGCTCTC CGCGCCCCAG CCGCCGGCTG CCAGCTTTTC GGGGCCCCGA 60
GTCGCACCCA GCGAAGAGAG CGGGCCCGGG ACAAGCTCGA ACTCCGGCCG CCTCGCCCTT 120
CCCCGGCTCC GCTCCCTCTG CCCCCCTCGGG GTCGCGGCC CACGATGCTG CAGGGCCCTG 180
GCTCGCTGCT GCTGCTCTTC CTGCGCTCGC ACTGCTGCCT GGGCTCGGCG CGCGGGCTCT 240
TCCTCTTTGG CCAGCCGAC TTCTCCTACA ACCGCAGMAA TTGCAAGCCC ATCCGGTCA 300
ACCTGCAGCT GTGCCACGGC ATCGAATACC AGAACATGCG GCTGCCAAC CTGCTGGCC 360
ACGAGACCAT GAAGGAGGTG CTGGAGCAGG CCGGGCGTTG GATCCCGCTG GTCATGAAGC 420
AGTGCCACCC GGACACCAAG AAGTTCCTGT GCTCGCTCTT CGCCCCCGTC TGCCCTCGATG 480
ACCTAGACGA GACCATCCAG CCATGCCACT CGCTCTGGGT GCAGGTGAAG GACCGCTGCG 540
CCCCGGTCAT GTCCGCCCTTC GGNTTCCCCT GGCCCGACAT GCTTGAGTGC GACCGTTTCC 600
CCCAAGGACAA CGACCTTTGC ATCCCCCTCG CTAGCAGCGA CCACCTCCTG CCAGCCACCG 660
AGGAAGCTCC AAAGGTATGT GAAGCCTGCA AAAATAAAA TGATGATGAC AACGACATAA 720
TGGAAACGCT TTGTAAAAAT GATTTTGAC TGAAAATAAA AGTGAAGGAG ATAACCTACA 780
TCAACCGAGA TACCAAAATC ATCCTGGAGA CCAAGAGCAA GACCATTAC AAGCTGAACC 840
GTGTGTCCGA AAGGGACCTG AAGAAATCGG TGCTGTGGCT CAAAGACAGC TTGCAGTGCA 900
CCTGTGAGGA GATGAACGAC ATCAACGCGC CCTATCTGGT CATGGGACAG AAACAGGGTG 960
GGGAGCTGGT GATCACCTCG GTGAAGCGGT GGCAGAAAGGG GCAGAGAGAG TTCAAGCGCA 1020
TCTCCCGCAG CATCCGCAAG CTGCAGTGCT AGTCCCAGCA TCCTGATGGC TCCGACAGGC 1080
CTGCTCCAGA GCACGGCTGA CCATTCTGC TCCGGGATCT CAGCTCCCGT TCCCCAAGCA 1140
CACTCCTAGC TGCTCCAGTC TCAGCCTGGG CAGCTTCCCC CTGCCTTTTG CACGTTTGCA 1200
TCCCCAGCAT TTCCGTAGTT ATAAGGCCAC AGGAGTGGAT AGCTGTTTC ACCTAAAGGA 1260

60

AAAGCCCACC CGAATCTTGT AGAAATATTC AACTAATAA AATCATGAAT ATTTTATGA 1320  
 AGTTAAAAAA TAGCTCACTT TAAAGCTAGT TTTGAATAGG TGCAACTGTG ACTTGGGCT 1380  
 5 GGTTGGTTGT TGTTTGTGT TTTGAGTCAG CTGATTTCA CTTCCCACTG AGGTTGTCA 1440  
 AACATGCAA TTGCTCAAT TTCTCTGTG GCCCAAACCTT GTGGGTCAACA ACCCTGTGT 1500  
 AGATAAAAGCT GGCTGTTATC TCAACATCTT CATCAGCTCC AGACTGAGAC TCAGTGTCTA 1560  
 10 AGCTTACAA CAATTCACTA TTTTATACCT TCAATGGAA CTTAAACTGT TACATGTATC 1620  
 ACATTCCAGC TACAATACTT CCATTATTA GAAGCACATT AACCATTCT ATAGCATGAT 1680  
 15 TTCTCAACT AAAAGGCAA AGATATAAT TTTATAATTG ACTTGAGTAC TTTAAGCCTT 1740  
 GTTTAAAACA TTTCTTACTT AACTTTGCA AATTAAACCC ATTGTAGCTT ACCTGTAATA 1800  
 20 TACATAGTAG TTTACCTTTA AAAGTGTAA AAATATTGCT TTAACCAACA CTGAAATAT 1860  
 TTCAGATAAA CATTATATTC TTGTATATAA ACTTTACATC CTGTTTACCC TAAAAAAA 1920  
 AAAAAAAAAA AAAAAACTCG AGGGGGGCC GGTACCCAAT 1960  
 25

## (2) INFORMATION FOR SEQ ID NO: 115:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 536 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GTCGTCAGCC CCCGGGCAC AGYAGGACGT TTGGGGCCT TCTTTCAGCA GGGGACAGCC 60  
 CGATTGGGA CAATGGGTC TCTTGGCAC ATCTTGGTTT TCTGTGTGGG TCTCCTCACC 120  
 ATGGCCAAGG CAGAAAGTCC AAAGGAACAC GACCCGTTCA CTTACGACTA CCAGTCCCTG 180  
 CAGATCGGAG GCCTCGTCAT CGCCGGATC CTCTTCATCC TGGGCATCCT CATCGTGTG 240  
 45 ACCAGAAAGAT GCGGGTGCAA GTTCAACCAG CAGCAGAGGA CTGGGAACC CGATGAAGAG 300  
 GAGGGAACCTT TCCGCAGCTC CATCCGCCGT CTGTCCAMCC GCANGCGTA GAAACACCTG 360  
 50 GAGCGATGGA ATCCGGCCAG GACTCCCCGT GCACCTGACA TCTCCCACGC TCCACCTGCG 420  
 CGCCCACCGC CCCCTCCGCC GCCCCCTCCC CAGCCCTGCC CCCGCAGACT CCCCTGCCG 480  
 CCAAGACTTC CAATAAAACG TGCGTTCTC TCGAMAAAAA AAAAAATAAA AAAACT 536  
 55

## (2) INFORMATION FOR SEQ ID NO: 116:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 790 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

	GTGGGGAGGG GGCGGAGCAA AGCCGCGCCT CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC	60
10	CTGACTTGAA CCTTCCCGGT CCCCAGCCCT CAACAGGAGG CGCAGAAAAT CTTCAAAGCC	120
	AACCACCCCA TGGACGCAGA AGTTACTAAG GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC	180
15	CTGGACAATG TGGACCCAA CCCTGAGAAC TTCGTGGGGG CGGGGATCAT CCAGACTAAA	240
	GCCCTGCAGG TGGGCTGTCT GCTTCGGCTG GAGCCAATG CCCAGGCCA GATGTACCGG	300
20	CTGACCCCTGC GCACCAGCAA GGAGCCCCGTC TCCCCTCACCC TGTGTGAGCT GCTGGCACAN	360
	AGTTCTGAGC CCTGGACTCT GCCCCGGGGG ATGTGGCCGG CACTGGCAG CCCCTTGGAC	420
	TGAGGCAGTT TTGGTGGATG GGGGACCTCC ACTGGTGACA GAGAAGACAC CAGGGTTTGG	480
25	GGGATGCCTG GGACTTTCTT CCGGCCTTTT GTATTTTAT TTTTGTTCAT CTGCTGCTGT	540
	TTACATTCTG GGGGGTAGG GGGAGTCCCC CTCCCTCCCT TTCCCCCCCCA AGCACAGAGG	600
	GGAGAGGGGC CAGGGAAGTG GATGTCTCCCT CCCCTCCAC CCCACCCCTGT TGTAGCCCC	660
30	CCTACCCCCCT CCCCATCCAG GGGCTGTGTA TTATTGTGAG CGAATAAACAA GAGAGACGTT	720
	AACAGCCCCA TGTCTGTGTC CATCACCCAN TGNTAGGTAG TCAAAGAACT GGGGTGAGGG	780
35	CATGCAGAGT	790

## 40 (2) INFORMATION FOR SEQ ID NO: 117:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 776 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

50	CAGCGCTGGA AGCAGCTGAG CCTGTGAGGG GTGGGGAGGG GGCGGAGCAA AGCCGCGCCT	60
	CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC CTGACTTGAA CCTTCCCGGT CCCCAGCCCT	120
	CAACAGGAGG CGCAGAAAAT CTTCAAAGCC AACCAACCCCA TGGACGCAGA AGTTACTAAG	180
55	GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC CTGGACAATG TGGACCCAA CCCTGAGAAC	240
	TTCGTGGGGG CGGGGATCAT CCAGACTAAA CCCCTGCAGG TGGGCTGTCT GCTTCGGCTG	300
60	GAGCCAATG CCCAGGCCA GATGTACCGG CTGACCCCTGC GCACCAACAA GGAGCCCCGTC	360

	TCCCGTCACC TGTGTGAGCT GCTGGCACAG AGTTCTGAGC CCTGGACTCT GCCCCGGGGG	420
5	ATGTGGCCGG CACTGGGCAG CCCCTTGGAC TGAGGCAGTT TTGGTGGATG GGGGACCTCC	480
	ACTGGTGACA GAGAAGACAC CAGGGTTTGG GGGATGCCG GGACTTTCCCT CCGGCCTTT	540
	GTATTTTAT TTTTGTTCAT CTGCTGCTGT TTACATTCTG GGGGGTTAGG GGGAGTCCCC	600
10	CTCCCTCCCT TTCCCCCCCAGCACAGAGG GGAGAGGGC CAGGGAACTG GATGTCCTCC	660
	CCCCTCCAC CCCACCCCTGT TGTAGCCCT CCTACCCCT CCCCATCCAG GGGCTGTGTA	720
15	TTATTGTGAG CGAATAAACAGAGACAGCN TAAAAAAA AAAAAAAAT TGAGGG	776

## (2) INFORMATION FOR SEQ ID NO: 118:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 453 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 25 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

30	GGTTCTGACA CCAGATGTC TCTGCTCCTG GTTAATGTCA GTGAGGGCTG GAAGTTGAAT	60
	AAATGAGAAC AGGAGTGGTC TGGGCCATG TAAATGATCC TCCCTTGAAA GGAGGAACAG	120
	CTTTCATCAT TTGTTCCAGC TAAGCCTTGC ATGCATTATA GATCTGGTGC TAAGCAGTGG	180
35	GAAAGATCTC ATAAGTAATG TTTTATGTC TTTCKGTC TCYTCCTCKG TTGTTCTTGG	240
	CTTGTGGGTT GTGTTTGKGG TTGTTAACTG GAAAATTGCT ATAAGCCAGT TGTCYCKAAK	300
	TTTWWAAAAC GAATTAGAAA AACCATAAAA TCYTCGGCC YATGCACATK GTCCCYGTT	360
40	TGTGAAAACA TTAAAGGTA AATAAAAAGG AAGGAGAACAGCAATAATG TGCATCAAAT	420
	ATATTCTGAG TTCTAGAGAA ATTAATGACC AAG	453

45

## (2) INFORMATION FOR SEQ ID NO: 119:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2016 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

55	AGGCTGTTCA CAGGCACCCC GAGACAGCGT CCCCCCTCTG GGCGCACTGG ATTTGACGTT	60
60	GCAGGACGCG CGGCTGGAAC CCCCAGGGCC CGCTGCTCAC AGACCGGGAC TCCGCCTCCG	120

	GTTCCCGAGG GCGTGGGAG GCGCTGCGGG ANCCCAACAG GATGCCTTCC GTGCCCTCCA	180
5	TCAAGATCTC AATTTTGTC GCAATTCTA CAGCCCCGT TGATTGGAGA GCTGGCTCCG GAAGAACCCA GCGAKGATGG ACCCCTGAAT GCGCATGGTC GAGGACTTCC GAGCCCTGCA	240
	CCAGGCAGCC GAGGACATGA AGCTGTTGA TGCCAGTCCC ACCTTCTTTG CTTCCTACT	300
10	GGGCCACATC CTGGCCATGG AGGTGCTGGC CTGGCTCCTT ATCTACCTCC TGGGTCTGG CTGGGTGCCG AGTGCCTCTGG NCCGCCTTCA TCCTGGCCAT CTCTCAGGCT CAGTCCTGGT	360
	GTCTGCAGCA TGACCTGGGC CATGCTCCAT CTTCAAGAAAG TCCTGGTGG ACCACGTGGC	420
15	CCAGAAGTTC GTGATGGGGC AGCTAAAGGG CTTCTCCGCC CACTGGTGG ACCACGTGGC CTTCCAGCAC CACGCCAAGC CCAACATCTT CCACAAAGAC CCAGACGTGA CGGTGGCGCC	480
	CGTYTTCTC CTGGGGAGT CATCCGTCGA GTATGGNCAA GAAGAAACGC AGATAACCTAC	540
20	CCTACAACCA GCAGCACCTG TACTTCTTCC TGATCGGCCG GCCGCTGTC ACCCTGGTGA ACTTTGAAGT GGAAAATCTG GCGTACATGC TGGTGTGCAT GCAGTGGCG GATTGCTCT	600
	GGGCGGCCAG CTTCTATGCC CGCTTCTTCT TATCCTACCT CCCCTCTAC GGCGTCCCTG GGGTGCTGCT CTTCTTTGTT GCTGTCAGGT ATGGCAGGGG GTGGCGAGGT CACACACAGG	660
25	CGACAGGTGA CCCCCACTGC AGCCCCCCAC CAGAGCTTCC CTTTCCCCTG CTGCAGAATG GGGCCAGTGG TACTGCCCTCC CTGGCTTGC GGTGGAATCA CATAAACACA AGYTTCAAGGA	720
	GCCCAGGGTC GGTGGGTITA GGGAGCGTGG CCTGGCTTGT AAGTGGCCCG GTGGGTGTCG GAGCTGCTCT GGACTCAGCC TCACAGTGG AACTGCTCCA TTCAGATTCT TTAAACACTG	780
30	GCAAGGGGGC GATGGCCACA ATCCTATTGT ACAGATAAGG AAGTCAAGGC CAYTTGGGG CAGYTGCTCT TCCAGCCTCC ACTCAGGGTG CCTTAAGTGG TGAGCTGGAC CTAGGGCAGT	840
	GGCGAGCYTC CCCACAGGGT CCTGGAAAGC CACTGGTTCG TGTGGATCAC ACAGATGAAC CACATCCCCA AGGAGATCGG CCACGAGAAG CACCGGGACT GGGTCAGCTC TCAGCTGGCA	900
35	GCCACCTGCA ACGTGGAGCC CTCACCTTTC ACCAACTGGT TCAGCGGGCA CCTCAACTTC CAGATCGAGC ACCACCTCTT CCCAGGATG CCGAGACACA ACTACAGCCG GGTGGCCCG	960
	CTGGTCAAGT CGCTGTGTGC CAAGCACGGC CTCAGCTACG AATGAAGCCC TTCCCTACCG CGCTGGTGGG CATCGTCAGG TCCCTGAAGA AGTCTGGTGA CATCTGGCTG GACGCCTACC	1020
40	TCCATCAGTG AAGGCAACAC CCAGGGGGC AGAGAAGGGC TCAGGGCACC AGCAACCAAG CCAGCCCCCG GCGGGATCGA TACCCCCAMC CCTCCACTGG CCAGCCTGGG GGTGCCCTGC	1080
	CTGCCCTCCT GGTACTGTTG TCTTCCCCCTC GGCCCCCTCA CATGTGTATT CAGCAGCCCT	1140
45	ATGGCCTTGG CTCTGGGCCT GATGGGACAG GGGTAGAGGG AAGGTGAGCA TAGCACATTT	1200
	CGATCGAGC ACCACCTCTT CCCAGGATG CCGAGACACA ACTACAGCCG GGTGGCCCG	1260
50	CTGGTCAAGT CGCTGTGTGC CAAGCACGGC CTCAGCTACG AATGAAGCCC TTCCCTACCG CGCTGGTGGG CATCGTCAGG TCCCTGAAGA AGTCTGGTGA CATCTGGCTG GACGCCTACC	1320
	TCCATCAGTG AAGGCAACAC CCAGGGGGC AGAGAAGGGC TCAGGGCACC AGCAACCAAG CCAGCCCCCG GCGGGATCGA TACCCCCAMC CCTCCACTGG CCAGCCTGGG GGTGCCCTGC	1380
55	CTGCCCTCCT GGTACTGTTG TCTTCCCCCTC GGCCCCCTCA CATGTGTATT CAGCAGCCCT	1440
	ATGGCCTTGG CTCTGGGCCT GATGGGACAG GGGTAGAGGG AAGGTGAGCA TAGCACATTT	1500
60	CGATCGAGC ACCACCTCTT CCCAGGATG CCGAGACACA ACTACAGCCG GGTGGCCCG	1560
	CTGGTCAAGT CGCTGTGTGC CAAGCACGGC CTCAGCTACG AATGAAGCCC TTCCCTACCG CGCTGGTGGG CATCGTCAGG TCCCTGAAGA AGTCTGGTGA CATCTGGCTG GACGCCTACC	1620
	TCCATCAGTG AAGGCAACAC CCAGGGGGC AGAGAAGGGC TCAGGGCACC AGCAACCAAG CCAGCCCCCG GCGGGATCGA TACCCCCAMC CCTCCACTGG CCAGCCTGGG GGTGCCCTGC	1680
	CTGCCCTCCT GGTACTGTTG TCTTCCCCCTC GGCCCCCTCA CATGTGTATT CAGCAGCCCT	1740
65	ATGGCCTTGG CTCTGGGCCT GATGGGACAG GGGTAGAGGG AAGGTGAGCA TAGCACATTT	1800
	CGATCGAGC ACCACCTCTT CCCAGGATG CCGAGACACA ACTACAGCCG GGTGGCCCG	1860
70	CTGGTCAAGT CGCTGTGTGC CAAGCACGGC CTCAGCTACG AATGAAGCCC TTCCCTACCG CGCTGGTGGG CATCGTCAGG TCCCTGAAGA AGTCTGGTGA CATCTGGCTG GACGCCTACC	1920

5  
 TCCTAGAGCG AGAATTGGGG GAAAGCTGTT ATTTTATAT TAAAATACAT TCAGATGTAA 1980  
 AAAAAAAA AAAAAANCT CGAGGGGGG CCCCCG 2016

10 (2) INFORMATION FOR SEQ ID NO: 120:

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2136 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

20 GGGGACGGAG CCGCTGTCAA CTCTCCAAT CAGCTCAGCT GATCGGTTGC CGCCGCCGCC 60  
 25 GCGCCAGAT TCTGGAGGCG AAGAACCAA AGCTGAGAAC ATGGACGTTA ATATGCC 120  
 ACTCCGCGCC TGGGACGATT TCTTCCCGG TTCCGATCGC TTTGCCCGGC CGGACTTCAG 180  
 30 GGACATTTCC AAATGGAACA ACCGCGTAGT GAGCAACCTG CTCTATTACC AGACCAACTA 240  
 CCTGGTGGTG GCTGCCATGA TGATTTCCAT TGTGGGTTT CTGAGTCCCT TCAACATGAT 300  
 CCTGGGAGGA ATCGTGGTGG TGCTGGTGT CACAGGGTTT GTGTGGGCAG CCCACAATAA 360  
 35 AGACGTCCCT CGCCGGATGA AGAACCGCTA CCCCACGACG TTCGTTATGG TGGTCATGTT 420  
 GGCAGCTAT TTCCATTATCT CCAATGTTGG AGGAGTCATG GTCTTTGTGT TTGGCATTAC 480  
 40 TTTTCCATTG CTGTTGATGT TTATCCATGC ATCGTTGAGA CTTCGGAACC TCAAGAACAA 540  
 ACTGGAGAAT AAAATGGAAG GAATAGGTTT GAAGAGGACA CCGATGGCA TTGTCCTGGA 600  
 TGCCTAGAA CAGCAGGAAG AAGGCATCAA CAGACTCACT GACTATATCA GCAAAGTGAA 660  
 45 GGAATAAACAA TAACCTACCT GAGCTAGGGT TGCAGCAGAA ATTGAGTTGC AGCTTGCCCT 720  
 TGTCCAGACC TATKTTCTGC TTGCGTTTTT GAAACAGGAG GTGCACGTAC CACCAATTAA 780  
 50 TCTATGGCAG CATGCATGTA TAGGCCAAC TATTATCAGC TCTGATGTTT CAGAGAGAAG 840  
 ACCTCAGAAA CCGAAAGAAA ACCACCAACCC TCCATTGTTG TCTGAAGTTT CACGTGTGTT 900  
 TATGAAATCT AATGGGAAAT GGATCACAGC ATTCTTTAA GGGATTAAA AAAAATAAAAA 960  
 55 GAATTACGGC TTTTACAGCA ACAATACGAT TATCTTATAG GAAAAAAA ATCAATTGAA 1020  
 AGTATCAAGA CAATACGAGT AAATGAAAAG GCTGTTAAAG TAGATGACAT CATGTGTAG 1080  
 CCTGTTCTA ATCCCCTAGA ATTGTAATGT GTGGGATATA AATTAGTTT TATTATTCTC 1140  
 TAAAAAATCA AAGATGATCT CTATCACTT GCCACCTGTT TGATGTGCAG TGGAAACTGG 1200  
 TTAAGCCAGT TGTTCATACT TCSTTTACAA ATATAAAGAT AGCTGTTAG GATATTGT 1260  
 60

	TACATTTTG TAAATTTTG AAATGCTAGT AATGTGTTT CACCAGCAAG TATTGTTGC	1320
	AAACTTAATG TCATTTCCCT TAAGATGGTT ACAGCTATGT AACCTGTATT ATTCTGGACG	1380
5	GACTTATTAA AATACAAACA GACAAAAAAT AAAACAAAAC TTGAGTTCTA TTTACCTTGC	1440
	ACATTTTTG TTGTTACAGT GAAAAAAATG GTCCAAGAAA ATGTTTGCCA TTTTTGCATT	1500
10	GTTCGTTTT TAACTGGAAC ATTTAGAAAG AAGGAAATGA ATGTGCATTT TATTAATTCC	1560
	TTAGGGGCAC AAGGAGGACA ATAATAGCTG ATCTTTGAA ATTTGAAAAA CGTCPTTAGA	1620
	TGACCAAGCA AAAAGACTTT AAAAAATGGT AATGAAAATG GAATGCAGCT ACTGCAGCTA	1680
15	ATAAAAATT TTAGATAGCA ATTGTTACAA CCATATGCCT TTATAGCTAG ACATTAGAAT	1740
	TATGATAGCA TGAGTTATA CATTCTATTA TTTTCCTCC CTTTCTCATG TTTTTATAAA	1800
	TAGGTAAATAA AAAATGTTT GCCTGCCAAT TGAATGATT CGTAGCTGAA GTAGAACAT	1860
20	TTAGGTTCT GTAGCATTAA ATTGTGAAGA CAACTGGAGT GGTACTTACT GAAGAAACTC	1920
	TCTGTATGTC CTAGAATAAG AAGCAATGAT GTGCTGCTTC TGATTTTCT TGCATTTAA	1980
25	ATTCTCAGCC AACCTACAGC CATGATCTT AGCACAGTGA TATCACCAGT ACTTCACAGA	2040
	CATGGTCTAG AATCTGTACC CTTACCCACA TATGAAGAAT AAAATTGATT AAAGGTTAAA	2100
	AAAAAAAAWAA AAAAMWAGG GGGGCCCGGT WCCCAAG	2136
30		

(2) INFORMATION FOR SEQ ID NO: 121:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

45	GGCCTAGTAT CTGGGCAGCT GTGCATGGAG ATAGCCAGAG GAAACATTCT TTTCTTAAT	60
	GRATGGTGA CCACATTTCG TTGTTCTTGC CTCCTATTAT CCGTGCSCTA TTTGCATSCCT	120
	GGTTTCTTCT ACAGTAGTTT ATGTAAATGT TGTTTTGTCC TTGTCGTTCT CAGTAGAATT	180
50	GGTTCTGTAA ACGAACCTG GTCTGTAAAT TTCAAGTATA	219

55 (2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1686 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

## (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

5	GCTGGAGATT CACATTTTAC CTGATTGCCT TCATTGCCGG CATGCCGTC ATITGGATA	60
	AACCCCTGGTT CTATGACATG AAGAAAGTTT GGGAGGGATA TCCCATAACAG AGCACTATCC	120
10	CTTCCCAGTA TTGGTACTAC ATGATTGAAC TTTCCTTCTA CTGGTCCCTG CTCTTCAGCA	180
	TTGCCCTCTGA TGTCAGCGA AAGGATTCA AGGAACAGAT CATCCACCAT GTGRCCACCA	240
	TCATTCTCAT CAGCTTTCC TGGTTGCCA ATTACATCCG AGCTGGGACT CTAATCATGG	300
15	CTCTGCATGA CTCTTCCGAT TACCTGCTGG AGTCAGCCAA GATGTTAAC TACGCCGGAT	360
	GGAAAGAACAC CTGCAACAAC ATCTTCATCG TCTTCGCCAT TGTTTTATC ATCACCCGAC	420
20	TGGTCATCCT GCCCTTCTGG ATCCTGCATT GCACCCCTGGT GTACCCACTG GAGCTCTATC	480
	CTGCCCTCTT TGGSTATTAC TTCTTCATT CCATGATGGG AGTTCTACAG CTGCTGCATA	540
	TCTTCTGGGC CTACCTCATT TTGCGCATGG CCCACAAGTT CATAACTGGG AAAGCTGGTA	600
25	GAAGATGAAC GCAWGCRGG GNAAGAAACA GAGAGCTAG AGGGGGAGGA GGCTGCAGCT	660
	GGGGGAGGAG CAAAGAGCCG GCCCCTAGCC AATGGCCACC CCATCCTCAA TAACAACCAT	720
30	CGTAAGAATG ACTGAACCCT TATTCCAGCT GCCTCCCAGA TTAATGCATA AACCCAAGGA	780
	ACTACCCYGC TCCCTGCGCT ATAGGGTCAC TTTAAGCTCT GGGGAAAAAG GAGAAAGTGA	840
	GAGGAGAGTT CTCTGCATCC TCCCTCCCTTG CTGTGACCC AGTTGCCTTT AAACCAAATT	900
35	CTAACCCAGCC TATCCCCAGG TAGGGGGACG TTGGTTATAT TCTGTTAGAG GGGGACGGTC	960
	GTATTTTCCT CCCTACCCCG CAAGTCATCC TTTCTACTGC TTTTGAGGCC CTCCCTCAGC	1020
40	TCTCTGTGGG TAGGGGTTAC AATTTCACATT CCTTATTCTG AGAATTGGC CCCAGCTGTT	1080
	TGCCTTTGAC TCCCTGACCT CCAGAGCCAG GGTTGTGCCT TATTGTCCCA TCTGTGGGCC	1140
	TCATTCTGCC AAAGCTGGAC CAAGGCTAAC CTTCTAAGC TCCCTAACTT GGGCCAGAAA	1200
45	CCAAAGCTGA GCTTTTAAC TCTCTCCCT ATGACACAAA TGAATTGAGG GTAGGAGGAG	1260
	GGTGCACATA ACCCTTACCC TACCTCTGCC AAAAAGTGGG GGCTGTACTG GGGACTGCTC	1320
50	GGATGATCCT TCTTAGTGCT ACTTCTTCA GCTGTCCCTG TAGCGACAGG TCTAAGATCT	1380
	GACTGCCTCC TCCCTTCTCT GGCCTCTTCC CCCTTCCCTC TTCTCTTCAG CTAGGCTAGC	1440
	TGGTTTGGAG TAGAATGGCA ACTAAATTCTA ATTTTATTT ATTAATATT TGGGGTTTGT	1500
55	GTTTTAAAGC CAGAATTACG GCTAGCACCT AGCAATTCAAG CAGAGGGACC ATTTTAGACC	1560
	AAAATGTACT GTTAATGGGT TTTTTTTAA ATTAAGA TAAATAAAA AATATTAANT	1620
60	AAAACATGGC AATAAGTGTC AGACTATTAG GAATTGAGAA GGGGATCAA CTAAATAAAC	1680

GAAGAG

1686

5

(2) INFORMATION FOR SEQ ID NO: 123:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1211 base pairs  
 10 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

15	CAGCCTGTGC CAGACGAGGA GGTGATTGAG CTGTATGGGG GTACCCAGCA CATCCCAC	60
	TACCAAGATGA GTGGCTTCTA TGGCAAGGGT CCCTCCATT AAGCAGTTCAT GGACATCTTC	120
20	TCCGCTACCGG AGATGGCTCT GCTGTCCGT GTGGTGGACT ACTTTCTGGG CCACAGCCTG	180
	GAGTTTGACC AAACATCTCT ACAAGGACGT GACGGACGCC ATCCGAGACG TGCATGTGAA	240
25	GGGCCTCATG TACCAAGTGGG TCGAGCAGGA CATGGAGAAG TACATCCTGA GAGGGGATGA	300
	GACGTTTGCT GTCCCTGAGCC GCCTGGTGGC CCATGGAAA CAGCTGTTCC TCATCACCAA	360
	CAGTCCTTTC AGCTTCGTAG ACAAGGGAT GCGGCACATG GTGGGTCCCC ATTGGCGCCA	420
30	CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCCAGCT TCTTCACTGA CCGGGCGAAC	480
	TTTCAGAAAA CTCGATGAGA AGGGCTCACT TCAGTGGGAC CGGATCACCC GCTTGGAAAA	540
	GGGCAAGATC TATCGGCAGG GAAACCTGTT TGACCTCTTA CGCTTGACGG AATGGCGTGG	600
35	CCCCCGCGTG CTCTACTTCG GGGACCACCT CTATAGTGAT CTGGCGGATC TCATGCTGCG	660
	GCACGGCTGG CGCACAGGCG CCATCATCCC CGAGCTGGAG CGTGAGATCC GCATCATCAA	720
40	CACGGAGCAG TACATGCACT CGCTGACGTG GCAGCAGGCG CTCACGGGC TGCTGGAGCG	780
	CATGCAGACC TATCAGGACG CGGAGTCGAG GCAGGTGCTG GCTGCCTGGA TGAAAGAGCG	840
	GCAGGGAGCTG AGGTGCATCA CCAAGGCCCT GTTCAATGCG CAGTTGGCA GCATCTTCCG	900
45	CACCTTCCAC AACCCCCACCT ACTTCTCAAG GCGCCTCGTG CGCTTCTCTG ACCTCTACAT	960
	GGCCTCCCTC AGCTGCCITGC TCAACTACCG CGTGGACTTC ACCTTCTACC CACGCCGTAC	1020
50	GCCGCTGCAG CACGAGGCAC CCCTCTGGAT GGACCAGCTT CTGCACCCGC TGCATGAAGA	1080
	CCCCCTTCCT TGGTGACATG GCCCACATCC GCTGAGGGCA CCTTTATTGT CTGGGACAGG	1140
	CCCTCAGCCC CTCCCTGCCCT ATCCACCCAG ACAAGCAATA AAAGTGGTCT CCTCCCTGAA	1200
55	AAAAAAAAAA A	1211

## (2) INFORMATION FOR SEQ ID NO: 124:

## (i) SEQUENCE CHARACTERISTICS:

- 5                   (A) LENGTH: 1804 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

10	CGCACCTATG GGCTCGCTAC CAGGACATGC GGAGACTGGT GCACGACCTC CTGCCCCCG	60
	AGGTCTGCAG TCTCCTGAAC CCAGCAGCCA TCTACGCCAA CAACGAGATC AGCCTGCGTG	120
15	ACGTGAGGT CTACGGCTTT GACTACGACT ACACCCCTGGC CCAGTATGCA GACGCACITGC	180
	ACCCCGAGAT CTTCA GTACAGTACC GCCCGTGACA TCCTGATCGA GCACATACAAG TACCCAGAAG	240
20	GGATTCGGAA GTATGACTAC AACCCCAGCT TTGCCATCCG TGCCCTCCAC TATGACATTG	300
	AGAAGAGCCT TCTGATGAAG ATTGACGCCT TCCACTACGT GCAGCTGGGG ACAGCCTACA	360
	GGGGCCTCCA GCCTGTGCCA GACGAGGAGG TGATTGAGCT GTATGGGGGT ACCCAGCACA	420
25	TCCCACATA CCAGATGAGT GGCTTCTATG GCAAGGGTCC CTCCATTAAG CAGTTCATGG	480
	ACATCTCTC GCTACGGAG ATGGCTCTGC TGTCCTGTGT GGTGGACTAC TTCTGGGCC	540
	ACAGCCTGGN AGTTTGACCA AGCACATCTC TACAAGGAGC TGACGGACGC CATCCGAGAC	600
30	GTGCATGTGA AGGGCCTCAT GTACCAGTGG ATCGAGCAGG ACATGGAGAA GTACATCCTG	660
	AGAGGGGATG AGACGTTTGC TGTCTTGAGC CGCCTGGTGG CCCATGGAA ACAGCTGTT	720
35	CTCATCACCA ACAGTCCTTT CAGCTTCGTA GACAAGGGGA TGCGGCACAT GGTGGGTCCC	780
	GATTGGGCC ACTCTTCGAT GTGGTCATTG TCCAGGCAGA CAAGCCCAGC TTCTTCACTG	840
40	ACCGCGCAA GCTTTTCAGA AAACTCGATG AGAAGGGCTC ACTTCAGTGG GACCGGATCA	900
	CCCGCTTGGAA AAAGGGCAAG ATCTATCGGC AGGGAAACCT GTTTGACTTC TTACGCTTGA	960
	CGGAATGGCG TGGCCCCCGC GTGCTCTACT TCGGGACCA CCTCTATAGT GATCTGGCG	1020
45	ATCTCATGCT CGGGCACGGC TGGCGCACAG GCGCCATCAT CCCCCAGCTG GAGCGTGAGA	1080
	TCCGCATCAT CAACACGGAG CAGTACATGC ACTCGCTGAC GTGGCAGCAG GCGCTCACGG	1140
50	GGCTGCTGGA GCGCATGCGAG ACCTATCAGG ACGCGGAGTC GAGGCAGGTG CTGGCTGCCT	1200
	GGATGAAAGA GCGGCAGGAG CTGAGGTGCA TCACCAAGGC CCTGTTCAAT GCGCAGTTG	1260
	GCAGCATCTT CGGCACCTTC CACAACCCA CCTACTCTC AAAGGCCTC CGTGCCTTC	1320
55	TCTGACCTCT ACATGGCTTC CCTCAGCTGC CTGCTCAACT ACCCGTGGA CTPCACCTTC	1380
	TACCCACGCC GTACGCCGCT GCAGCAGGAG GCACCCCTCT GGATGGACCA GCTCTGGACC	1440
60	GGCTGCATGA AGACCCCTT CCTGGGTGAC ATGGCCCACA TCCGCTGAGG GCACCTTAT	1500

	TGTCTGGGAC AGGCCCTCAG CCCCTCCTGC CCCATCCACC CAGACAAGCA ATAAAAGTGG	1560
	TCTCCTCCCT GTGCATGCTT CTGCTTCAG CCCCAGCCTC GTCACTTGAC TGTGAGGATC	1620
5	CTCTGGGTGT CAGGGAAGTC CTCCCTCCAGC AGTGAGTCAT CGAAGGGTTC ACAAAAGGTG	1680
	TCGCTGCCAA AGACAGGGTT GGGGACAGAG ACCAGGGTGG GGTTGGTCCC TTCTTGCCAC	1740
10	GGTGAGAAGT CGTCGTCAGC CGGACGGGTG GGTCGACCCG GGAATTCCGG ACCGGTACCT	1800
	GCAG	1804

15

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

20	(A) LENGTH: 1282 base pairs
	(B) TYPE: nucleic acid
	(C) STRANDEDNESS: double
	(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

	CCGCAGGNCA GCGACGCGAC TCTGGTGCAG GCCGCTTCT TCCCCCGAG CTGGGCGTGC	60
	GCGGCCGCAA TGAACTGGGA GCTGCTGCTG TGGCTGCTGG TGCTGTGCGC GCTGCTCCTG	120
30	CTCTTGGTGC AGCTGCTGCG CTTCTGAGG GCTGACGGCG ACCTGACGCT ACTATGGGCC	180
	GAGTGGCAGG GACGACGCC AGAATGGAG CTGACTGATA TGGTGGTGTG GGTGACTGGA	240
	GCCTCGAGTG GAATTGGTGA GGAGCTGGCT TACCAAGTTGT CTAAACTAGG AGTTTCTCTT	300
35	GTGCTGTCAG CCAGAAGAGT GCATGAGCTG GAAAGGGTGA AAAGAAGATG CCTAGAGAAT	360
	GGCAATTAA AAGAAAAAGA TATACTTGTT TTGCCCCCTTG ACCTGACCGA CACTGGTCC	420
40	CATGAAGCGG CTACCAAAGC TGTCTCCAG GAGTTGGTA GAATCGACAT TCTGGTCAAC	480
	AATGGTGGAA TGTCCCAGCG TTCTCTGTC ATGGATACCA GCTTGGATGT CTACAGAAAG	540
45	CTAATAGAGC TTAACTACTT AGGGACGGTG TCCTTGACAA AATGTGTTCT GCCTCACATG	600
	ATCGAGAGGA AGCAAGGAAA GATTGTTACT GTGAATAGCA TCCTGGGTAT CATATCTGTA	660
	CCTCTTCCA TTGGATACTG TGCTAGCAAG CATGCTCTCC GGGGTTTTTT TAATGGCCTT	720
50	CGAACAGAAC TTGCCACATA CCCAGGTATA ATAGTTCTA ACATTTGCCG AGGACCTGTG	780
	CAATCAAATA TTGTGGAGAA TTCCCTAGCT GGAGAAGTCA CAAAGACTAT AGGCAATAAT	840
55	GGAGACCAAGT CCCACAAGAT GACAACCAAGT CGTTGTGTGC GGCTGATGTT AACACCATG	900
	GCCAATGATT TGAAAGAAAGT TTGGATCTCA GAACAAACCTT TCTTGTAGT AACATATTG	960
	TGGCAATACA TGCCAACCTG GGCCTGGTGG ATAACCAACA AGATGGGAA GAAAAGGATT	1020
60	GAGAACTTTA AGAGTGGTGT GGATGCAGAC TCTTCTTATT TTAAAATCTT TAAGACAAAA	1080

270

	CATGACTGAA AAGAGCAYCT GTACTTTCA AGCCACTGGA GGGARAATG GAAAACATGA	1140
5	AAACAGCAAT CTTCTTATGC TTCTGAATAA TCAAAGACTA ATTTGTGRIT TTACTTTTA	1200
	ATAGATATGA CTTTGCTTCC AACATGGAAT GAAATAAAA ATAATAATA AAAGATTGCC	1260
	ATGGAAAAAA AAAAGNNNGG AN	1282

10

## (2) INFORMATION FOR SEQ ID NO: 126:

## 15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1296 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

GGCAGAGCTT AGAGTGTGGA AAAGGCAACC AGGTTGGCCG TAAGTGCCTG CTGGAATGCG	60
25 TGTGCCTCCA CASGGRTCTG GGCATCCGGA CTGATAACCA GCCGCCAGA CTGAGGGATG	120
GAAGGCAGTG AGATGGGGC CCGTCCAGGC GGACACCCGC AGAAATGGAG CTTCTGTGG	180
30 TCTCTTGAC TCTGGCTGCC TCTTGCCTC TCTGTGTCTC TCTTCTTGG TCTCTCCCTC	240
TCTCCTCCCTC AGCCTGGTCT TTCTCTTGG TGACACACTTA GTTATTGTGG TGAGCAATGG	300
AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GGACACAGCC TAAAAGGAC	360
35 AAAAAGTTAG AAGACAGCAT AGCAACTCAG CTCAGGRGC TACCAAGAGAA AAATAGCAAC	420
TGATGTGGGT GCTTTTTTTT TTTTTTTAAT TTGAATAAAA AGAATTAGAA GTGATGTCCT	480
40 TTTATAAAAT GCCTTCTCCC CCTTCCGCC TACAGTCTCT TCCTCTCCCC TTAGAGGGGG	540
GAAAGTGTAT AAACCTACAG GGTTGTGAGT CTGAAAAGAG GATCCCCCTC ACCCCCACCC	600
TGGGCAGAGC AGTGGGGTTT GGGGGTGCG AGAGGGGAC ACAGATCCTG GCACACTGTG	660
45 GATATTCTT GCAGATTGCA GTCTCTGTG GCCCACAG AGTGGTAGA CTATGCCTC	720
TGGCAGGTGC CACCTTTGG TACCAACATG TTCTGAGGTG TTAGGATTTG GGTGGGTTT	780
50 TTTTGTGTTG TTTTTTTTTT CCNTTTGGTC TTTTTTTTT TCYCCCTKTA AAGAAAAGCT	840
AAAGGCCGCT GTGAGTCCTG GTGGCAGGCT CTCCATGGAT GTACCATATC GAAGATAATT	900
TTTATACTGC ATTATTTATGG ATTATTTGT AATGTGTGAT TCCGTCTGCT GAGGAGGTGG	960
55 GAGGGGCTCC AGGGAAAGCC ACCCACCTTC AGTGAGGTTG CTCCCCAGCT GAGGCCACCG	1020
GGCATGGGAT GTGGAGGCTG GCGACACACC CTGTCCTCT CCAAGGCTGG GCGCGTGGGG	1080
CGTCCAGAGT CTCTCTGGGT CTCAGATGTC CATCTGCCAC CTCTTGTAA GGCTCTAGCC	1140

60

AGAAGGGAGG GTGAGGGTAG AAGAAAGTTA TTCCCGAAGA AAAAAAGAAT GAAAAGTCAT	1200
TGTACTGAAC TGTGTTTATA TTTTAAAG TTACTATTAA AAGCGGACGT CGTGGGTCGA	1260
<b>5 CCCGGGAATT CCCGGACCAGG TACTGTCAAG TCTAAC</b>	<b>1296</b>

**10 (2) INFORMATION FOR SEQ ID NO: 127:**

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 737 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

20 GGCANAGTGG AGGCAATGCC AGCTCCAGGA CAGAGGCTCA GGTGCCAAC GGGCAAGGCA	60
GCCCAGGGGG CTGTGTCGT TCAAGTCAGG CTTCCCCGGC CCYTCGGCA NCAGCGCTTC	120
25 CACGGGCAGC CCGGGGCCCC ACCCCACGCA CTGAAGAGGC CGCCTGGCT GCCATGGCCC	180
TGACCTTCCT GCTGGTGCTG CTCACCTGG CCACGCTCTG CACACGGCTG CACAGAACT	240
TCCGACGCGG GGAGAGCATC TACTGGGGC CCACAGCGGA CAGCCAGGAC ACAGTGGCTG	300
30 CTGTGCTGAA GOGGAGGCTG CTGCAGCCCT CGCGCCGGGT CAAGCGCTCG CGCCGGAGAC	360
CCYTCYTCCC GCCCACGCCG GACAGCGGCC CGGAAGGCGA GAGCTGGAG TGACGGCTG	420
GGACCTGCCA CTGTGGCGTG CGGTCTCCCC GCGCCGCGAG GCGCGAMCT NTGCCACGTG	480
35 GACCGCGCGC NGGGCGCTMC CCTGGTGGCG ATGGCGCGGC ACTGGCGAGC ACTGCGKGGG	540
CTTTCCTCCT TGTTGGTTGC TGAGTGGGCG GCCAAGGGGA GAAAAGGAGC CGCTTYTGCC	600
40 TCCCTTGCCA AAACCTCCGTT TCTAATTAAA TTATTTTTAG TAGAAAAAAA AAAAAAAAAA	660
AAAAAAAAAA AAAAAAAAAC TCGAGGGGGG GCGCGGTACC CAATTNGCCA	720
AATAGCGATC GTATNAA	737

**45**

**50 (2) INFORMATION FOR SEQ ID NO: 128:**

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1925 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

60 CCCCCGCTCC AAAGCTAACCC CTGGGGCTTG AGGGGAAGAR GCTGACTGTA CGTTCCTTCT	60
---	----

	ACTCTGGCAC CACTCTCCAG GCTGCCATGG GGCCCAGCAC CCCTCTCCTC ATCTTGTTC	120
	TTTTGTCAAG GTCGGGACCC CTCCAAGGAC AGCAGCACCA CCTTGTGGAG TACATGGAAC	180
5	GCCGACTAGC TGCTTTAGAG GAACGGCTGG CCCAGTGCCA GGACCAGAGT AGTCGGCATG	240
	CTGCTGAGCT GCGGGACTTC AAGAACAAAGA TGCTGCCACT GCTGGACGTG CGAGAGAAGG	300
10	AGCGGGAGGC ACTCAGAACT GAGGCCGACA CCATCTCCGG GAGAGTGGAT CGTCTGGAGC	360
	GGGAGGTTAGA CTATCTGGAG ACCCAGAACCC CAGCTCTGCC CTGTGTAGAG TTGATGAGA	420
	AGGTGACTGG AGGCCCTGGG ACCAAAGGCA AGGAAGAAG GAATGAGAAG TAGATATGG	480
15	TGACAGACTG TGGCTACACA ATCTCTCAAG TGAGATCAAT GAAGATTCTG AAGCGATTG	540
	GTGGCCCAGC TGGTCTATGG ACCAAGGATC CACTGGGCA AACAGAGAAG ATCTACGTGT	600
20	TAGATGGGAC ACAGAAATGAC ACAGCCTTG TCTTCCAAG GCTGCGTGC TTCACCCCTG	660
	CCATGGCTGC CCGGAAAGCT TCCCGAGTCC GGGTGCCTT CCCCTGGTA GGCACAGGGC	720
	AGCTGGTATA TGGTGGCTTT CTTTTATTTTG CTGGGAGGCC TCCTGGAAGA CCTGGTGGAG	780
25	GTGGTGAGAT GGAGAACACT TTGCGACTAA TCAAATTCCA CCTGGCAAAC CGAACAGTGG	840
	TGGACAGCTC AGTATTCCCA GCAGAGGGC TGATCCCCC CTACGGCTTG ACAGCAGACA	900
30	CCTACATCGA CCTGGCAGCT GATGAGGAAG GTCTTGGGC TGTCTATGCC ACCCGGGAGG	960
	ATGACAGGCA CTTGTGTCTG GCCAAGTTAG ATCCACAGAC ACTGGACACA GAGCAGCAGT	1020
	GGGACACACC ATGTCCCAGA GAGAATGCTG AGGCTGCCCT TKTCATCTGT GGGACCCCT	1080
35	ATGTCGTCTA TAACACCGT CCTGCCAGTC GGGCCCGCAT CCAGTGCTCC TTTGATGCCA	1140
	GCGGACCCCTG ACCCTGAAC GGGCAGCACT CCCTTATTTT CCCCCAGAT ATGGTGCCTA	1200
40	TGCCAGCCTC CGCTATAACC CCCGAGAACG CCAGCTCTAT GCCTGGGATG ATGGCTACCA	1260
	GATTGTCTAT AAGCTGGAGA TGAGGAAGAA AGAGGAGGAG GTTGTAGGAG CTAGCCTTGT	1320
	TTTTTGATC TTTCTCACTC CCATACATTT ATATTATATC CCCACTAAAT TTCTTGTTC	1380
45	TCATTCTTCA AATGTGGCC AGTTGTGGCT CAAATCCTCT ATATTTTAG CCAATGGCAA	1440
	TCAAATTCTT TCAGCTCCCTT TGTTCATAC GGAACCTCCAG ATCCTGAGTA ATCCTTTAG	1500
50	AGCCCGAAGA GTCAAAACCC TCAATGTTCC CTCTGCTCT CCTGCCCAT GTCAACAAAT	1560
	TTCAGGCTAA GGATGCCCA GACCCAGGGC TCTAACCTTG TATGCGGGCA GGCCCAGGGA	1620
	GCAGGGCAGCA GTGTTCTTCC CCTCAGAGTG ACTTGGGAG GGAGAAATAG GAGGAGACGT	1680
55	CCAGCTCTGT CCTCTCTTCC TCACTCCTCC CTTCAGTGTGCTGAGGAACA GGACTTTCTC	1740
	CACATTGTT TGTATTGCAA CATTGAT TAAAAGGAAA ATCCAMAAA AAAAAAAA	1800
60	AAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA	1860

ACTGCGGCCG CTGTCCCTTC TGTCGTCTTC TCGCAGCCGT ACCCTTCTGT CGTCTTCTCG	1920
CAGCC	1925

5

## (2) INFORMATION FOR SEQ ID NO: 129:

- 10           (i) SEQUENCE CHARACTERISTICS:  
               (A) LENGTH: 2713 base pairs  
               (B) TYPE: nucleic acid  
               (C) STRANDEDNESS: double  
               (D) TOPOLOGY: linear

15           (xii) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

TCTTACCTTC CCAACCCCTCT GGCATCCCCA GCACGTGATGG TCCTGGCATC CACGGCTGAG	60
20       GCCAGCCGTG ACTGCTTCCA TCCCTTGTCGA GCAGGCCACGA CCCTTTGGTG TACCTGTYTC	120
AGTTGACAAG GACGTGCATA TTCCCTTCAC CAACGGTTCC TATACCTTTG CCTCTATGTA	180
25       CCATCGGCAA GGTGGGGTGC CAGGCACCTTG TGCCAATCGT GATTTCCCCC CTTCTCTACT	240
ACACCTCCAC CCTCAATTG CTCCCCAAA TCTAGATTGC ACCCCAATCA GTATGCTGAA	300
TCATAAGTGG TGTGGGGGTT TCCGGCCATT GSCTCCACCC GRGGACCGGG RGAGYTATCA	360
30       GTCAGCTTTA CGCCGGCCAA GCGACTTAAG AACTGCCATG ACACAGAGTC TCCCCACTTG	420
CGCNCTCAG ATGCAGATGG GAANGAATAT GACTTTGGGA CACAGCTGCM ATCTAGCTCC	480
CCCGGTTCAC TAAAGGTGGA TGACACTGGG AAGAAGATTG TTGCTGTCTC TGGCCTCATT	540
35       TCTGATCGGG AAGCCTCATC TAGCCCAGAG GNTCGGNAAT GACAGATGTA AGAAGAAAGC	600
AGCGGCATTG TTGACAGAGCC AGGCCCAAT TTGCCCCATC TGCCAGGTCC TGCTGAGGCC	660
40       CAGTGAGCTG CAGGAGCATA TGGAGCAGGA ACTGGAGCAG CTAGCCCAAC TGCCCTCGAG	720
CAAGAATTCC CTTCTGAAGG ATGCCATGGC TCCAGGCACC CCAAAGTCCC TCCTGTTGTC	780
TGCTTCCATC AAGAGGGAAG GAGAGTCTCC AACGGCATCA CCCCACTCAT CTGCCACCGA	840
45       TGACCTCCAC CATTCAAGACA GATACCAAGAC CTTTCTGCGA GTACGAGCCA ACCGGCAGAC	900
CCGAYTGAAT GYTCGGATTG CGAAAATGAA ACGGAGGAAG CAAGATGAAG GGCAGGTATG	960
50       TCCCCTGTGC AACCGCCCCC TGGCAGGATC GGAGCAGGAG ATGAGTAGGC ATGTGGAGCA	1020
TTGCCTTCT AAGAGGGAAG GCTCCTGCAT GGCTGAGGAT GATGCTGTGG ACATCGAGCA	1080
TGAGAACAAAC AACCGCTTTG AGGAGTATGA GTGGTGTGGA CAGAAGCGGA TACGGGCCAC	1140
55       CACTCTCCTG GAAGGTGGCT TCCGAGGCTC TGGCTTCATC ATGTGCAGCG GCAAAGAGAA	1200
CCCGGACAGT GATGCTGACT TGGATGTGGA TGGGGATGAC ACTCTGGAGT ATGGGAAGCC	1260
60       ACAATACACA GAGGCTGATG TCATCCCTG CACAGGCCAG GAGCCTGGTG AAGCCAAGGA	1320

	GAGAGAGGCA CTTGGGGCG CAGTCCTAAA TGGCGGCCCT CCCAGCACGC GCATCACACC	1380
5	TGAGTTCTCT AAATGGCCA GTGATGAGAT GCCATCCACC AGCAATGGTG AAAGCAGCAA GCAGGAGGCC ATGCAGAAGA CCTGCAAGAA CAGGGACATC GAGAAAATCA CCGAAGATTG	1440 1500
	AGCTGTGACC ACGTTTGAGG CTCTGAAGGC TCGGGTCAGA GAACCTGAAC GGCAGCTATC	1560
10	TCGTGGGAC CGTTACAAAT GCCTCATCTG CATGGACTCG TACTCGATGC CCCTAACGTC CATCCAGTGT TGGCACGTGC ACTGCGAGGA GTGCTGGCTG CGGACCCCTGG GTGCCAAGAA	1620 1680
	GCTCTGCCCT CAGTGAACA CGATCACAGC GCCCGGAGAC CTGCGGAGGA TCTACTTGIG	1740
15	AGCTATCTGC CCCAGGCAGG CCTCGCCTCC AGCAGCCCCA CCTGCCCTCA GCCTCTGTGA CAGTGACCGT YTCCCTTTGT ACATACTTGC ACACAGGTTG CCCATGTACA TACATGCACA	1800 1860
20	TACTCAAACA TGGGTACACA CACACACATT TACACACGC GGAECTCTGGA GCCAGAGTAG AGGCTGTGGC CCAGGGACTA CCTGCTGGCT CCCACCTATG GTTGGGGGC CATACTGTG	1920 1980
	CCAGCTCTGT TCCCACGGTG GGGCAGGGAG GTGGGGGTG GGGGAGTAGT GGGGCACGGC	2040
25	TCCTAAGATC CAGCCCCAT ACTGACAGAC GGACAGACAG ACATGCAAAC ACCAGACTGA AGCACATGTA ATATAGACCG TGTATGTTA CAATGTTGIG TATAATGGG ACAACTCTC	2100 2160
30	GCCCTCTACC TGTCCCCCTCC CCCTTTGGTT GTATGATTTT CTCTTTTTT AAGAACCCCT GGAAGCAGCG CCTCCCTTCAG GGTGGCTGG GAGCTCGGCC CATCCACCTC TTGGGGTAYC	2220 2280
	TGCCTCTCTC TCTCCCTGTGG TGTCCTTCC CTCTCCCATG TGCTCGGTGT TCAGTGGTGT	2340
35	ATATTTCTTC TCCCAGACAT GGGGCACACG CCCCAAGGGG CATGATCTC TCCCTAGTCT TAGCTCATGG GGCTCTTTAT AAGGAGTTGG GGGGTAGAGG CAGGAAATGG GAACCGAGCT	2400 2460
40	GAAGCAGAGG CTGAGTTAGG GGGCTAGAGG ACAGTGTCTC TGGCCACCCA GCCTCTGCTG AGAACCATTC CTGGGATTAG AGCTGCCTTT CCCAGGGAAA AAGTGTGTC TCCCCGACCC	2520 2580
	TCCCGTGGGC CCTGTGGTGT GATGCTGTGT CTGTATATTG TATACAAAGG TACTTGTCT	2640
45	TTCCCTTTGT AACTACATT TGACATGGAT TAAACCAGTA TAAACAGTTA AAAAAAAA AAAAAAAAACT CGA	2700 2713
50		

## (2) INFORMATION FOR SEQ ID NO: 130:

- 55           (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1011 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

	AGAGGACGGT GTGACCCGGG AGGAAGTAGA GCCTGAGGAG GCTGAAGAAC GCATCTCTGA	60
5	GCAACCCCTGC CCAGCTGACA CAGAGGTGGT GGAAGACTCC TTGAGGCAGC GTAAAAGTCA	120
	GCATGCTGAC AAGGGACTGT AGATTTAATG ATGCCTTTTC AAGAAATACAC ACCAAAACAA	180
10	TATGTCAGCT TCCCCTTGCG CTGCAGTTG TACCAAATCC TTAATTGAGC YTGAATGAGC	240
	AAGCTTCTCT TAAAAGATGC TCTCTAGTCA TTTGGTCTCA TGGCAGTAAG CCTCATGTAT	300
	ACTAAGGAGA GTCTTCCAGG TGTGACAATC AGGATATAGA AAAACAAACG TAGTGTNTGG	360
15	GATCTGTTG GAGACTGGGA TGGGAACAAG TTCATTTACT TAGGGGTAG AGAGTCTCGA	420
	CCAGAGGAGG CCATTCCCAG TCCTAATCAG CACCTTCCAG AGACAAGGCT GCAGGCCCTG	480
	TGAAATGAAA GCCAAGCAGG AGCCTGGCT CTGAGNCATC CCCAAAGTGT AACGTAGAAG	540
20	CCTTGCATCC TTTTCTTGTG TAAAGTATTT ATTTTGTCA AATTGCAGGA AACATCAGGC	600
	ACCACAGTGC ATGAAAAATC TTTCACAGCT AGAAATTGAA AGGGCCTTGG GTATAGAGAG	660
25	CAGCTCAGAA GTCATCCCAG CCCTCTGAAT CTCTGTGCT ATGTTTTATT TCTTACCTTT	720
	AATTTTCCA GCATTTCCAC CATGGCATT CAGGCTCTCC ACACITCTICA CTATTATCTC	780
	TTGGTCAGAG GACTCCAATA ACAGCCAGGT TTACATGAAC TGTGTTGTT CATTCTGACC	840
30	TAAGGGTTT AGATAATCAG TAACCATAAC CCCGTGAAGCT GTGACTGCCA AACATCTCAA	900
	ATGAAATGTT GTRGCCATCA GAGACTCAAA AGGAAGTAAG GATTITACAA GACAGATTAA	960
35	AAAAAAATIG TTTTGTCCAA AAAANAAAAA AAAAAAAACTC GAAGGGGGGG C	1011

## 40 (2) INFORMATION FOR SEQ ID NO: 131:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2278 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

50	GTAATTCGGC ACGAGGCGCC CAACATGGCG GGTGGCCGCT GCGGCCCGCA SCTAACGGCG	60
	CTCCTGGCCG CCTGGATCGC GGCTGTGGCG GCGACGGCAG GCCCCGAGGA GGCGCGCTG	120
	CCGCCGGAGC AGAGCCGGGT CCAGCCCAG ACCGCCCTCCA ACTGGACGCT GGTGATGGAG	180
55	GGCGAGTGGGA TGCTGAAATT TTACGCCCA TGGTGTCCAT CCTGCCAGCA GACTGATTCA	240
	GAATGGGAGG CTTTGCAAA GAATGGTGAATACATCAGA TCAGTGTGGG GAAGGTAGAT	300
60	GTCATTCAAG AACCAAGGTTT GAGTGGCCGC TTCTTGTCA CCACTCTCCC AGCATTTTT	360

	CATGCAAAGG ATGGGATATT CGCCCGTTAT CGTGGCCCCAG GAATCTTCGA AGACCTGCAG	420
	AATTATATCT TAGAGAAGAA ATGGCAATCA GTCGAGCCTC TGACTGGCTG GAAATCCCCG	480
5	GCTTCTCTAA CGATGTCTGG AATGGCTGGT CTTTTAGCA TCTCTGGCAA GATATGGCAT	540
	CTTCACAAC TTTTACAGT GACTCTTGGA ATTCTGCTT GGTGTTCTTA TGTCTTTTC	600
10	GTCATAGCCA CCTTGGTTT TGGCCTTTT ATGGGCTGG TCTTGGTGGT AATATCAGAA	660
	TGTTTCTATG TGCCACTTCC AAGGCATTTA TCTGAGCGTT CTGAGCAGAA TCGGAGATCA	720
	GAGGAGGCTC ATAGAGCTGA ACAGTTGCAG GATGCGGAGG AGGAAAAAGA TGATTCAAAT	780
15	GAAGAAGAAA ACAAAAGACAG CCTTGTAGAT GATGAAGAAG AGAAAGAAGA TCTTGGCGAT	840
	GAGGATGAAG CAGAGGAAGA AGAGGAGGAG GACAACTTGG CTGCTGGTGT GGATGAGGAG	900
20	AGAAAGTGAGG CCAATGATCA GGGGCCCA GGAGAGGACG GTGTGACCCG GGAGGNAAGT	960
	AGAGCCTGAG GAGGCTGAAG AAGGCATCTC TGAGCAACCC TGCCCAGCTG ACACAGAGGT	1020
	GGTGGAAAGAC TCCTTGAGGC AGCGTAAAG TCAGCATGCT GNCAAGGGAC TGTAGATTAA	1080
25	ATGATGCGTT TTCAAGAATA CACACAAAAA CAATATGTCA GCTTCCCTTT GGCTGCACT	1140
	TTGTACCAAA TCCCTAACCTT TTCCCTGAATG AGCAAGCTTC TCTTAAAGA TGCTCTCTAG	1200
30	TCATTTGGTC TCATGGCACT AAGCCTCATG TATACTAAGG AGAGTCTTCC AGGTGTGACA	1260
	ATCAGGATAT AGAAAAACAA ACGTAGTGTN TGGGATCTGT TTGGAGACTG GGATGGGAAC	1320
	AAGTTCAATT ACTTAGGGGT CAGAGAGTCT CGACCAGAGG AGGCCATTCC CAGTCCTAAT	1380
35	CAGCACCTTC CAGAGACAAG GCTGCAGGCC TGTGAATGA AAGCCAAGCA GGAGCCTTGG	1440
	CTCTGAGGCA TCCCCAAAGT GTAACGTAGA AGCCTTGCAT CCTTTCTTGT TGTAAAGTAT	1500
40	TTATTTTGT CAAATTGCAG GAAACATCAG GCACCCAGT GCATGAAAAA TCTTTCACAG	1560
	CTAGAAATTG AAAGGGCCTT GGGTATAGAG AGCAGCTCAG AAGTCATCCC AGCCCTCTGA	1620
	ATCTCCCTGTG CTATGTTTA TTTCTTACCT TTAATTTTC CAGCATTCC ACCATGGGCA	1680
45	TTCAGGCTCT CCACACTCTT CACTATTATC TCTTGGTCAG AGGACTCCAA TAACAGCCAG	1740
	GTTTACATGA ACTGTGTTTG TTCATTCTGA CCTAAGGGGT TTAGATAATC AGTAACCATA	1800
50	ACCCCTGAAG CTGTGACTGC CAAACATCTC AAATGAAATG TTGTRGCCAT CAGAGACTCA	1860
	AAAGGAAGTA AGGATTTAC AAGACAGATT AAAAAAAAT TGTGTTGTCC NAAAATATAG	1920
	TTGTTGTTGA TTTTTTTTA AGTTTCTAA GCAATTTTT TCAAGCCAGA AGTCCTCTAA	1980
55	GTCTTGCCAG TACAAGGTAG TCTTGTGAAG AAAAGTTGAA TACTGTTTG TTTTCATCTC	2040
	AAGGGGTTCC CTGGGTCTTG AACTACTTTA ATAATAACTA AAAAACCACT TCTGATTTTC	2100
60	CTTCAGTGAT GTGCTTTGG TGAAAGAATT AATGAACCTCC AGTACCTGAA AGTGAAGAT	2160

TTGATTTGT TTCCATCTTC TGAAATCTTC CAAAGAATTAA TATCTTGTAA AATCTCTCAA 5 TACTCAATCT ACTGTAAGTA CCCAGGGRRGG STAATTTCYT TAAAAAAAAA AAAAAAAA	2220 2278
---	--------------

## 10 (2) INFORMATION FOR SEQ ID NO: 132:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1088 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - 15 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

GGCAGGGCG GCGTGAACCC GTCGGCACT GTGTCCTGA CAATGGAAC AGCCGACAGT 20 GATGAGATGG CCCCGGAGCC CCACAGCACA CCCACATCGA TGTGCACATC CACCAGGAGT CTGCCCTGGC CAAGCTCCTG CTCACCTGCT GCTCTGCGCT GCGGCCCCGG GCCACCCAGG 25 CCAGGGGCAG CANCCGGCTG CTGGTGGCCT CGTGGGTGAT GCAGATCGTG CTGGGGATCT TGAGTGCAGT CCTAGGAGGA TTTTTCTACA TCCCGGACTA CACCCCTCTC GTCACCTCGG 30 GAGCTGCCAT CTGGACAGGG GCTGTGGCTG TGCTGGCTGG AGCTGCTGCC TTCATTTAYG AGAAACGGGG TGGTACATAC TGGGCCCTGC TGAGGACTCT GCTARCGCTG GCAGCTTTCT 35 CCACAGCCAT CGCTGCCCTC AAACTTTGGAA ATGAAGATTT CCGATATGGC TACTCTTATT ACAACAGTGC CTGGCCGCATC TCCAGCTCGA GTGACTGGAA CACTCCAGCC CCCACTCAGA 40 GTCCAGAAGA AGTCAGAAGG CTACACCTAT GTACCTCCCT CATGGACATG CTGAAGGCCT TGTTCAGAAC CCTTCAGGCC ATGCTCTTGG GTGTCCTGGAT TCTGCTGCC TTGGCATCTC TGGCCCTCT GTGGCTGTAC TGCTGGAGAA TGTTCCCAAC CAAAGGGAAA AGAGACCAGA 45 AGGAAATGTT GGAAGTGAGT GGAATCTAGC CATGCCCTCTC CTGATTATTA GTGCCTGGTG CTTCTGCACC GGGCGTCCCT GCATCTGACT GCTGGAAGAA GAACCAGACT GAGGAAAAGA 50 GGCTCTTCAA CAGCCCCAGT TATCCTGGCC CCATGACCGT GGCCACAGCC CTGCTCCAGC AGCACTTGCC CAITCCTTAC ACCCCCTTCCC CATCCTGCTC CGCTTCATGT CCCCTCCTGA 55 GTAGTCATGT GATAATAAAC TCTCATGTTA TTGTTCCNAA AAAAAAAAAA AAAAAAAAAT TGGGGGGGGG CCGGTACCCA TTGGGCCTNN GGGGGNGGTT TAAAATTAAT GGGGGGGGTT TAAAAGGG 1088	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080
--	---

## 60 (2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 553 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

10	GGCAGAGAGC AGATGGCCTT GACACCAGCA GGGTGACATC CGCTATTGCT ACTTCCTCTGC TCCCCCACAG TTCCCTCTGGA CTTCTCTGGA CCACAGTCCT CTGCCAGACC CCTGCCAGAC	60 120
15	CCCAAGTCCAC CATGATCCAT CTGGGTACA TCCTCTTCCT GCCTTTGCTC CCAGTGGCTG CAGCTCAGAC GACTCCAGGA GAGAGATCAT CACTCCCTGC CTTTTACCCCT GGCACTTCAG	180 240
20	GCTCTTGTTC CGGATGTGGG TCCCTCTCTC TGCCGCTCCT GGCAGGCCTC GTGGCTGCTG ATGCGGTGGC ATCGCTGCTC ATCGTGGGGG CGGTGTTCTCT GTGCGCACGC CCACGCCGA	300 360
25	GCCCCGCCA AGATGGCAA GTCTACATCA ACATGCCAGG CAGGGGCTGA CCCTCTGCA GCTTGGACCT TTGACTCTG ACCCTCTCAT CCTGGATGGT GTGTGGTGGC ACAGGAACCC CCGCCCCAAC TTTTGGATTG TAATAAAACA ATTGAAACAC CAAAAAAA AAAAAAAA	420 480 540
	AAAAAAAAAA AAA	553

30

(2) INFORMATION FOR SEQ ID NO: 134:

**35** (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 467 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

40 Met Arg Pro Gln Glu Leu Pro Arg Leu Ala Phe Pro Leu Leu Leu Leu  
1 5 10 15

45 Leu Leu Leu Leu Pro Pro Pro Pro Cys Pro Ala His Ser Ala Thr  
                   20                   25                   30

Arg Phe Asp Pro Thr Trp Glu Ser Leu Asp Ala Arg Gln Leu Pro Ala  
35 40 45

50 Trp Phe Asp Gln Ala Lys Phe Gly Ile Phe Ile His Trp Gly Val Phe  
50 55 60

55 Glu Lys Ile Pro Lys Tyr Val Glu Phe Met Lys Asp Asn Tyr Pro Pro  
85 90 95

60 Xaa Phe Lys Tyr Glu Asp Phe Gly Pro Leu Phe Thr Ala Lys Phe Phe  
100 105 110

	Asn Ala Asn Gln Trp Ala Xaa Ile Phe Gln Ala Ser Gly Ala Lys Tyr		
	115	120	125
5	Ile Val Leu Thr Ser Lys His His Glu Gly Phe Thr Leu Trp Gly Ser		
	130	135	140
	Glu Tyr Ser Trp Asn Trp Asn Ala Ile Asp Glu Gly Pro Lys Arg Asp		
	145	150	155
10	Ile Val Lys Glu Leu Glu Val Ala Ile Arg Asn Arg Thr Asp Leu Arg		
	165	170	175
	Phe Gly Leu Tyr Tyr Ser Leu Phe Glu Trp Phe His Pro Leu Phe Leu		
15	180	185	190
	Glu Asp Glu Ser Ser Ser Phe His Lys Arg Gln Phe Pro Val Ser Lys		
	195	200	205
20	Thr Leu Pro Glu Leu Tyr Glu Leu Val Asn Asn Tyr Gln Pro Glu Val		
	210	215	220
	Leu Trp Ser Asp Gly Asp Gly Gly Ala Pro Asp Gln Tyr Trp Asn Xaa		
	225	230	235
25	Thr Gly Phe Leu Ala Trp Leu Tyr Asn Glu Ser Pro Val Arg Gly Thr		
	245	250	255
	Val Val Thr Asn Asp Arg Trp Gly Ala Gly Ser Ile Cys Lys His Gly		
30	260	265	270
	Gly Phe Tyr Thr Cys Ser Asp Arg Tyr Asn Pro Gly His Leu Leu Pro		
	275	280	285
35	His Lys Trp Glu Asn Cys Met Thr Ile Asp Lys Leu Ser Trp Gly Tyr		
	290	295	300
	Arg Arg Glu Ala Gly Ile Ser Asp Tyr Leu Thr Ile Glu Glu Leu Val		
	305	310	315
40	Lys Gln Leu Val Glu Thr Val Ser Cys Gly Gly Asn Leu Leu Met Asn		
	325	330	335
	Ile Gly Pro Thr Leu Asp Gly Thr Ile Ser Val Val Phe Glu Glu Arg		
45	340	345	350
	Leu Arg Gln Met Gly Ser Trp Leu Lys Val Asn Gly Glu Ala Ile Tyr		
	355	360	365
50	Glu Thr His Thr Trp Arg Ser Gln Asn Asp Thr Val Thr Pro Asp Val		
	370	375	380
	Trp Tyr Thr Ser Lys Pro Lys Glu Lys Leu Val Tyr Ala Ile Phe Leu		
	385	390	395
55	Lys Trp Pro Thr Ser Gly Gln Leu Phe Leu Gly His Pro Lys Ala Ile		
	405	410	415
	Leu Gly Ala Thr Glu Val Lys Leu Leu Gly His Gly Gln Pro Leu Asn		
60	420	425	430

Trp Ile Ser Leu Glu Gln Asn Gly Ile Met Val Glu Leu Pro Gln Leu  
435 440 445

5 Thr Ile His Gln Met Pro Cys Lys Trp Gly Trp Ala Leu Ala Leu Thr  
450 455 460

Asn Val Ile  
465

10

(2) INFORMATION FOR SEQ ID NO: 135:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 222 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

20 Met Trp Ser Ala Gly Arg Gly Gly Ala Ala Trp Pro Val Leu Leu Gly  
1 5 10 15

Leu Leu Leu Ala Leu Leu Val Pro Gly Gly Gly Ala Ala Lys Thr Gly  
25 20 25 30

Ala Glu Leu Val Thr Cys Gly Ser Val Leu Lys Leu Asn Thr His  
35 40 45

30 His Arg Val Arg Leu His Ser His Asp Ile Lys Tyr Gly Ser Gly Ser  
50 55 60

Gly Gln Gln Ser Val Thr Gly Val Glu Ala Ser Asp Asp Ala Asn Ser  
65 70 75 80

35 Tyr Trp Arg Ile Arg Gly Gly Ser Glu Gly Gly Cys Arg Arg Gly Ser  
85 90 95

40 Pro Val Arg Cys Gly Gln Ala Val Arg Leu Thr His Val Leu Thr Gly  
100 105 110

Lys Asn Leu His Thr His His Phe Pro Ser Pro Leu Ser Asn Asn Gln  
115 120 125

45 Glu Val Ser Ala Phe Gly Glu Asp Gly Glu Asp Asp Leu Asp Leu  
130 135 140

Trp Thr Val Arg Cys Ser Gly Gln His Trp Glu Arg Glu Ala Ala Val  
145 150 155 160

50 Arg Phe Gln His Val Gly Thr Ser Val Phe Leu Ser Val Thr Gly Glu  
165 170 175

Gln Tyr Gly Ser Pro Ile Arg Gly Gln His Glu Val His Gly Met Pro  
55 180 185 190

Ser Ala Asn Thr His Asn Thr Trp Lys Ala Met Glu Gly Ile Phe Ile  
195 200 205

60 Lys Pro Ser Val Glu Pro Ser Ala Gly His Asp Glu Leu Xaa

210

215

220

## 5 (2) INFORMATION FOR SEQ ID NO: 136:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 156 amino acids
- (B) TYPE: amino acid

- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

Met	Val	Ile	Glu	Ile	Ser	Asn	Lys	Thr	Ser	Ser	Ser	Thr	Cys	Ile
1														15

15

Leu	Val	Leu	Leu	Val	Ser	Phe	Cys	Leu	Leu	Leu	Val	Pro	Ala	Met	Tyr
															30
20								25							

15

Ser	Ser	Asp	Thr	Arg	Gly	Ser	Leu	Pro	Ala	Glu	His	Gly	Val	Leu	Ser
															45
20							35		40						

20

Arg	Gln	Leu	Arg	Ala	Leu	Pro	Ser	Glu	Asp	Pro	Tyr	Gln	Leu	Glu	Leu
															55
50								55				60			

25

Pro	Ala	Leu	Gln	Ser	Glu	Val	Pro	Lys	Asp	Ser	Thr	His	Gln	Trp	Leu
															80
65						70			75						

Asp	Gly	Ser	Asp	Cys	Val	Leu	Gln	Ala	Pro	Gly	Asn	Thr	Ser	Cys	Leu
															95
85							90								

30

Leu	His	Tyr	Met	Pro	Gln	Ala	Pro	Ser	Ala	Glu	Pro	Pro	Leu	Glu	Trp
															110
100							105								

35

Pro	Phe	Pro	Asp	Leu	Phe	Ser	Glu	Pro	Leu	Cys	Arg	Gly	Pro	Ile	Leu
															125
115						120									

Pro	Leu	Gln	Ala	Asn	Leu	Thr	Arg	Lys	Gly	Gly	Trp	Leu	Pro	Thr	Gly
															140
130						135									

40

Ser	Pro	Ser	Val	Ile	Leu	Gln	Asp	Arg	Tyr	Ser	Gly				
145					150				155						

45

## (2) INFORMATION FOR SEQ ID NO: 137:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid

- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

Met	Met	Ile	Leu	Phe	Asn	Leu	Leu	Ile	Phe	Leu	Cys	Gly	Ala	Ala	Leu
1								5							15

55

Leu	Ala	Val	Gly	Ile	Trp	Val	Ser	Ile	Asp	Gly	Ala	Ser	Phe	Leu	Lys
															30
20							25								

60

Ile	Phe	Gly	Pro	Leu	Ser	Ser	Ser	Ala	Met	Gln	Phe	Val	Asn	Val	Gly
															45
35						40									

Tyr Phe Leu Ile Ala Ala Gly Val Val Val Phe Ala Leu Gly Phe Leu  
 50 55 60

5 Gly Cys Tyr Gly Ala Lys Thr Glu Ser Lys Cys Ala Leu Val Thr Phe  
 65 70 75 80

Phe Phe Ile Leu Leu Ile Phe Ile Ala Glu Val Ala Ala Val  
 85 90 95

10 Val Ala Leu Val Tyr Thr Thr Met Ala Glu His Phe Leu Thr Leu Leu  
 100 105 110

Val Val Pro Ala Ile Lys Lys Asp Tyr Gly Ser Gin Glu Asp Phe Thr  
 15 115 120 125

Gln Val Trp Asn Thr Thr Met Lys Gly Leu Lys Cys Cys Gly Phe Thr  
 130 135 140

20 Asn Tyr Thr Asp Phe Glu Asp Ser Pro Tyr Phe Lys Glu Asn Ser Ala  
 145 150 155 160

Phe Pro Pro Phe Cys Cys Asn Asp Asn Val Thr Asn Thr Ala Asn Glu  
 165 170 175

25 Thr Cys Thr Lys Gln Lys Ala His Asp Gln Lys Val Glu Gly Cys Phe  
 180 185 190

Asn Gln Leu Leu Tyr Asp Ile Arg Thr Asn Ala Val Thr Val Gly Gly  
 30 195 200 205

Val Ala Ala Gly Ile Gly Gly Leu Glu Leu Ala Ala Met Ile Val Ser  
 210 215 220

35 Met Tyr Leu Tyr Cys Asn Leu Gln Xaa  
 225 230

## 40 (2) INFORMATION FOR SEQ ID NO: 138:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

Met Gly Ser Ser Arg Trp Ser Val Ala Cys Pro Thr Gly Leu Gly Val  
 1 5 10 15

45 50 Leu Met Leu Gly Leu Gly Gly Asp His Pro Pro Gly Ser Gln Val Asp  
 20 25 30

Pro Leu Leu Met Gly Xaa Cys Val Arg Pro Xaa Leu Pro Glu Leu Thr  
 55 35 40 45

Ala Xaa Trp Arg Glu Xaa Gln Xaa Arg Ser Ala Ser Ala  
 50 55 60

## (2) INFORMATION FOR SEQ ID NO: 139:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 73 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

10	Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser Phe Ser			
	1	5	10	15
	Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys Thr Arg			
	20		25	30
15	Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met Gly Trp			
	35	40	45	
20	Gly Asp Trp Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala Asn Leu			
	50	55	60	
	Asp Lys Met Ala Ser Glu Gly Met Xaa			
	65	70		

25

## (2) INFORMATION FOR SEQ ID NO: 140:

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 377 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

35	Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Leu Met			
	1	5	10	15
	Gln Phe Leu Cys His Glu Phe Leu Arg Gly Asn Pro Arg Val Thr Arg			
	20		25	30
40	Leu Leu Ser Glu Met Arg Ile His Leu Leu Pro Ser Met Asn Pro Asp			
	35	40	45	
45	Gly Tyr Glu Ile Ala Tyr His Arg Gly Ser Glu Leu Val Gly Trp Ala			
	50	55	60	
	Glu Gly Arg Trp Asn Asn Gln Ser Ile Asp Leu Asn His Asn Phe Ala			
	65	70	75	80
50	Asp Leu Asn Thr Pro Leu Trp Glu Ala Gln Asp Asp Gly Lys Val Pro			
	85		90	95
	His Ile Val Pro Asn His Leu Pro Leu Pro Thr Tyr Tyr Thr Leu			
	100		105	110
55	Pro Asn Ala Thr Val Ala Pro Glu Thr Arg Ala Val Ile Lys Trp Met			
	115	120	125	
60	Lys Arg Ile Pro Phe Val Leu Ser Ala Asn Leu His Gly Gly Glu Leu			
	130	135	140	

Val Val Ser Tyr Pro Phe Asp Met Thr Arg Thr Pro Trp Ala Ala Arg  
145 150 155 160

5 Glu Leu Thr Pro Thr Pro Asp Asp Ala Val Phe Arg Trp Leu Ser Thr  
165 170 175

Val Tyr Ala Gly Ser Asn Leu Ala Met Gln Asp Thr Ser Arg Arg Pro  
180 185 190

10 Cys His Ser Gln Asp Phe Ser Val His Gly Asn Ile Ile Asn Gly Ala  
195 200 205

Asp Trp His Thr Val Pro Gly Ser Met Asn Asp Phe Ser Tyr Leu His  
15 210 215 220

Thr Asn Cys Phe Glu Val Thr Val Glu Leu Ser Cys Asp Lys Phe Pro  
225 230 235 240

20 His Glu Asn Glu Leu Pro Gln Glu Trp Glu Asn Asn Lys Asp Ala Leu  
245 250 255

Leu Thr Tyr Leu Glu Gln Val Arg Met Gly Ile Ala Gly Val Val Arg  
260 265 270

25 Asp Lys Asp Thr Glu Leu Gly Ile Ala Asp Ala Val Ile Ala Val Asp  
275 280 285

Gly Ile Asn His Asp Val Thr Thr Ala Trp Gly Asp Tyr Trp Arg  
30 290 295 300

Leu Leu Thr Pro Gly Asp Tyr Met Val Thr Ala Ser Ala Glu Gly Tyr  
305 310 315 320

35 His Ser Val Thr Arg Asn Cys Arg Val Thr Phe Glu Glu Gly Pro Phe  
325 330 335

Pro Cys Asn Phe Val Leu Thr Lys Thr Pro Lys Gln Arg Leu Arg Glu  
340 345 350

40 Leu Leu Ala Ala Gly Ala Lys Val Pro Pro Asp Leu Arg Arg Arg Leu  
355 360 365

Glu Arg Leu Arg Gly Gln Lys Asp Xaa  
45 370 375

## (2) INFORMATION FOR SEQ ID NO: 141:

50

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

Met Ile Cys Leu Ile Leu Leu Gln Ala Val Val Phe Leu Arg Ser  
1 5 10 15

60 Leu His Val Val His Asn Phe Gln Ile Leu Asp Leu Ser Gly Thr Ser

20                    25                    30

25

30

Tyr Pro Lys Phe Tyr Gln Thr Leu His Arg Gln  
35 40

5

(2) INFORMATION FOR SEQ ID NO: 142:

**10 (i) SEQUENCE CHARACTERISTICS:**

- (A) LENGTH: 41 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

15 Met Val His Val Leu Glu Ile Leu Leu Phe Ile Thr Met Gln Ala Val  
1 5 10 15

20	Ser Phe Pro Phe Gln Thr Gln Ile Asp Thr Cys Asn Thr Gln Asp Pro
	20                            25                            30
35	Ala Glu Arg Gln Pro Ala Ser Ile Val
	35                            40

25

(2) INFORMATION FOR SEQ ID NO: 143:

30                   (i) SEQUENCE CHARACTERISTICS:  
                      (A) LENGTH: 70 amino acids  
                      (B) TYPE: amino acid  
                      (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

35 Met Gly Ser Cys Ser Lys Asn Arg Ser Phe Phe Trp Met Thr Gly Leu  
1 5 10 15

Leu Val Phe Ile Ser Leu Leu Leu Ser Glu Trp Gln Gly Pro Trp Glu  
 20                    25                    30

40 Gly Arg Ala Ile Gly Glu Gly Trp Ala Ser Trp Ala Leu Thr Asn Gly  
35 40 45

45 Trp Ala Val Gln Leu Leu Met Ser Leu Gly Asn Asn Thr Glu Lys His  
           50                       55                       60

Ser Val Met Ile Tyr Glu  
65 70

50

(2) INFORMATION FOR SEQ ID NO: 144:

55        (i) SEQUENCE CHARACTERISTICS:  
                (A) LENGTH: 483 amino acids  
                (B) TYPE: amino acid  
                (D) TOPOLOGY: linear  
  
60        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:  
  
60        Met Ala Thr Gly Gly Gly Ile Arg Ala Met Thr Ser Leu Tyr Gly Gln

1	5	10	15	
Leu Ala Gly Leu Lys Glu Leu Gly Leu Leu Asp Cys Xaa Ser Tyr Ile				
	20	25	30	
5	Thr Gly Ala Ser Gly Ser Thr Trp Ala Leu Ala Asn Leu Tyr Lys Asp			
	35	40	45	
10	Pro Glu Trp Ser Gln Lys Asp Leu Ala Gly Pro Thr Glu Leu Leu Lys			
	50	55	60	
Thr Gln Val Thr Lys Asn Lys Leu Gly Val Leu Ala Pro Ser Gln Leu				
	65	70	75	80
15	Gln Arg Tyr Arg Gln Glu Leu Ala Glu Arg Ala Arg Leu Gly Tyr Pro			
	85	90	95	
Ser Cys Phe Thr Asn Leu Trp Ala Leu Ile Asn Glu Ala Leu Leu His				
	100	105	110	
20	Asp Glu Pro His Asp His Lys Leu Ser Asp Gln Arg Glu Ala Leu Ser			
	115	120	125	
25	His Gly Gln Asn Pro Leu Pro Ile Tyr Cys Ala Leu Asn Thr Lys Gly			
	130	135	140	
Gln Ser Leu Thr Thr Phe Glu Phe Gly Glu Trp Cys Glu Phe Ser Pro				
	145	150	155	160
30	Tyr Glu Val Gly Phe Pro Lys Tyr Gly Ala Phe Ile Pro Ser Glu Leu			
	165	170	175	
Phe Gly Ser Glu Phe Phe Met Gly Gln Leu Met Lys Arg Leu Pro Glu				
	180	185	190	
35	Ser Arg Ile Cys Phe Leu Glu Gly Ile Trp Ser Asn Leu Tyr Ala Ala			
	195	200	205	
40	Asn Leu Gln Asp Ser Leu Tyr Trp Ala Ser Glu Pro Ser Gln Phe Trp			
	210	215	220	
Asp Arg Trp Val Arg Asn Gln Ala Asn Leu Asp Lys Glu Gln Val Pro				
	225	230	235	240
45	Leu Leu Lys Ile Glu Glu Pro Pro Ser Thr Ala Gly Arg Ile Ala Glu			
	245	250	255	
Phe Phe Thr Asp Leu Leu Thr Trp Arg Pro Leu Ala Gln Ala Thr His				
	260	265	270	
50	Asn Phe Leu Arg Gly Leu His Phe His Lys Asp Tyr Phe Gln His Pro			
	275	280	285	
55	His Phe Ser Thr Trp Lys Ala Thr Thr Leu Asp Gly Leu Pro Asn Gln			
	290	295	300	
Leu Thr Pro Ser Glu Pro His Leu Cys Leu Leu Asp Val Gly Tyr Leu				
	305	310	315	320
60	Ile Asn Thr Ser Cys Leu Pro Leu Leu Gln Pro Thr Arg Asp Val Asp			

287

	325	330	335
	Leu Ile Leu Ser Leu Asp Tyr Asn Leu His Gly Ala Phe Gln Gln Leu		
	340	345	350
5	Gln Leu Leu Gly Arg Phe Cys Gln Glu Gln Gly Ile Pro Phe Pro Pro		
	355	360	365
	Ile Ser Pro Ser Pro Glu Glu Gln Leu Gln Pro Arg Glu Cys His Thr		
10	370	375	380
	Phe Ser Asp Pro Thr Cys Pro Gly Ala Pro Ala Val Leu His Phe Pro		
	385	390	395
15	Leu Val Ser Asp Ser Phe Arg Glu Tyr Ser Ala Pro Gly Val Arg Arg		
	405	410	415
	Thr Pro Glu Glu Ala Ala Ala Gly Glu Val Asn Leu Ser Ser Ser Asp		
	420	425	430
20	Ser Pro Tyr His Tyr Thr Lys Val Thr Tyr Ser Gln Glu Asp Val Asp		
	435	440	445
	Lys Leu Leu His Leu Thr His Tyr Asn Val Cys Asn Asn Gln Glu Gln		
25	450	455	460
	Leu Leu Glu Ala Leu Arg Gln Ala Val Gln Arg Arg Arg Gln Arg Arg		
	465	470	475
30	Pro His Xaa		

35 (2) INFORMATION FOR SEQ ID NO: 145:

	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 226 amino acids		
	(B) TYPE: amino acid		
40	(D) TOPOLOGY: linear		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:		
	Met Glu Gly Ala Pro Pro Gly Ser Leu Ala Leu Arg Leu Leu Phe		
	1	5	10
	15		
45	Val Ala Leu Pro Ala Ser Gly Trp Leu Thr Thr Gly Ala Pro Glu Pro		
	20	25	30
50	Pro Pro Leu Ser Gly Ala Pro Gln Asp Gly Ile Arg Ile Asn Val Thr		
	35	40	45
	Thr Leu Lys Asp Asp Gly Asp Ile Ser Lys Gln Gln Val Val Leu Asn		
	50	55	60
55	Ile Thr Tyr Glu Ser Gly Gln Val Tyr Val Asn Asp Leu Pro Val Asn		
	65	70	75
			80
	Ser Gly Val Thr Arg Ile Ser Cys Gln Thr Leu Ile Val Lys Asn Glu		
	85	90	95
60			

Asn Leu Glu Asn Leu Glu Glu Lys Glu Tyr Phe Gly Ile Val Ser Val  
 100 105 110  
 Arg Ile Leu Val His Glu Trp Pro Met Thr Ser Gly Ser Ser Leu Gln  
 5 115 120 125  
 Leu Ile Val Ile Gln Glu Glu Val Val Glu Ile Asp Gly Lys Gln Val  
 130 135 140  
 10 Gln Gln Lys Asp Val Thr Glu Ile Asp Ile Leu Val Lys Asn Arg Gly  
 145 150 155 160  
 Val Leu Arg His Ser Asn Tyr Thr Leu Pro Leu Glu Glu Ser Met Leu  
 165 170 175  
 15 Tyr Ser Ile Ser Arg Asp Ser Asp Ile Leu Phe Thr Leu Pro Asn Leu  
 180 185 190  
 Ser Lys Lys Glu Ser Val Ser Ser Leu Gln Thr Thr Ser Gln Tyr Leu  
 20 195 200 205  
 Ile Arg Asn Val Glu Thr Thr Val Asp Glu Asp Val Leu Pro Gly Gln  
 210 215 220  
 25 Val Thr  
 225

30 (2) INFORMATION FOR SEQ ID NO: 146:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 45 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

Met Gly Met Gly Ala Phe Gln Ala Phe Phe Trp Val Ile Leu Thr Val  
 1 5 10 15  
 40 Ser Asn Val Cys Val Leu Phe Lys Met Ser Leu Phe Phe Leu Leu Thr  
 20 25 30  
 Leu Ile Ser Lys Leu His Gly Asp Ala Glu Val Cys Xaa  
 45 35 40 45

50 (2) INFORMATION FOR SEQ ID NO: 147:

- 50 (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 132 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

Met Ser Gly Gly Trp Met Ala Gln Val Gly Ala Trp Arg Thr Gly Ala  
 1 5 10 15  
 60 Leu Gly Leu Ala Leu Leu Leu Leu Gly Leu Gly Leu Gly Leu Glu

	20	25	30
	Ala Pro Arg Ala Arg Phe Pro Pro Arg Pro Leu Pro Arg Pro His Pro		
	35	40	45
5	Ser Ser Gly Ser Cys Pro Pro Thr Lys Phe Gln Cys Arg Thr Ser Gly		
	50	55	60
	Leu Cys Val Pro Leu Thr Trp Arg Cys Asp Arg Thr Trp Thr Ala Ala		
10	65	70	75
	Met Ala Ala Met Arg Arg Ser Ala Gly Leu Ser His Val Pro Arg Lys		
	85	90	95
15	Gly Asn Ala His Arg Pro Leu Ala Ser Pro Ala Pro Ala Pro Ala Ser		
	100	105	110
	Val Thr Ala Leu Gly Glu Leu Thr Arg Asn Cys Ala Thr Ala Ala Ala		
	115	120	125
20	Trp Pro Ala Xaa		
	130		

25

(2) INFORMATION FOR SEQ ID NO: 148:

	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 92 amino acids		
30	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:		
	Met Glu Ala Thr Leu Glu Gln His Leu Glu Asp Thr Met Lys Asn Pro		
35	1	5	10
	15		
	Ser Ile Val Gly Val Leu Cys Thr Asp Ser Gln Gly Leu Asn Leu Gly		
	20	25	30
40	Cys Arg Gly Thr Leu Ser Asp Glu His Ala Gly Val Ile Ser Val Leu		
	35	40	45
	Ala Gln Gln Ala Ala Lys Leu Thr Ser Asp Pro Thr Asp Ile Pro Val		
	50	55	60
45	Val Cys Leu Glu Ser Asp Asn Gly Asn Ile Met Ile Gln Lys His Asp		
	65	70	75
	80		
50	Gly Ile Thr Val Ala Val His Lys Met Ala Ser Xaa		
	85	90	

55

(2) INFORMATION FOR SEQ ID NO: 149:

60

	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 165 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:		

Met Glu Pro Leu Arg Leu Leu Ile Leu Leu Phe Val Thr Glu Leu Ser  
 1 5 10 15

5 Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu  
 20 25 30

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys  
 35 40 45

10 Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val  
 50 55 60

Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly  
 15 65 70 75 80

Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr  
 85 90 95

20 Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser  
 100 105 110

Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val  
 115 120 125

25 Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro  
 130 135 140

Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser  
 30 145 150 155 160

Arg Ser Ser Ser Xaa  
 165

35

## (2) INFORMATION FOR SEQ ID NO: 150:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 139 amino acids  
 (B) TYPE: amino acid  
 (C) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

45 Met Ile Ser Leu Thr Asp Thr Gln Lys Ile Gly Met Gly Leu Thr Gly  
 1 5 10 15

Phe Gly Val Phe Phe Leu Phe Phe Gly Met Ile Leu Phe Phe Asp Lys  
 20 25 30

50 Ala Leu Leu Ala Ile Gly Asn Val Leu Phe Val Ala Gly Leu Ala Phe  
 35 40 45

Val Ile Gly Leu Glu Arg Thr Phe Arg Phe Phe Gln Lys His Lys  
 55 50 55 60

Met Lys Ala Thr Gly Phe Phe Leu Gly Gly Val Phe Val Val Leu Ile  
 65 70 75 80

60 Gly Trp Pro Leu Ile Gly Met Ile Phe Glu Ile Tyr Gly Phe Phe Leu

	85	90	95
	Leu Phe Arg Gly Phe Phe Pro Val Val Val Gly Phe Ile Arg Arg Val		
	100	105	110
5	Pro Val Leu Gly Ser Leu Leu Asn Leu Pro Gly Ile Arg Ser Phe Val		
	115	120	125
	Asp Lys Val Gly Glu Ser Asn Asn Met Val Xaa		
10	130	135	

(2) INFORMATION FOR SEQ ID NO: 151:

15	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 58 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:		
	Met Ser Ala Pro Gln Thr Arg Ile Ser Arg Ala Leu Val Leu Phe		
	1	5	10
			15
25	Leu Ala Pro Thr Leu Leu Ser Leu Gly His Gly Ile His Pro Ile Asn		
	20	25	30
	Thr Ala Thr Pro Tyr Xaa Thr Asp Gln Ala Lys Leu Ala Pro Gly Thr		
	35	40	45
30	Lys Glu Leu Asn His Asp Gln Ser Val Thr		
	50	55	

(2) INFORMATION FOR SEQ ID NO: 152:

35	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 48 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:		
	Met Ile Arg Lys Leu His Lys Ile Ile Val Phe Ser Pro Arg Val Ile		
	1	5	10
			15
	Val Leu Leu Asn Cys Phe Phe Ile Lys Ala Lys Phe Val Leu Tyr		
	20	25	30
50	Ile Phe Val Phe His Val Leu Asp Gly Ser Ile Ser Tyr Pro Val Xaa		
	35	40	45

55

(2) INFORMATION FOR SEQ ID NO: 153:

60	(i) SEQUENCE CHARACTERISTICS:		
----	-------------------------------	--	--

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

5

Met Leu Leu Asn Gln His Phe Lys Ile Phe Gly Ser Leu Ile His Met  
1 5 10 15

10

Asn Leu Leu Phe Ala Leu Ile Ser Leu Gly Ser Ser Asn Leu Ser Gly  
20 25 30

15

Val Gln Phe Cys Cys Glu Thr Val Gln Xaa  
35 40

(2)

INFORMATION FOR SEQ ID NO: 154:

(i) SEQUENCE CHARACTERISTICS:

20

(A) LENGTH: 72 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

25

Met Leu Ser Leu Ser Phe Leu Leu Arg Arg Val Leu Phe Leu Gly Phe  
1 5 10 15

30

Leu Gln Ala Ser Val Gly Glu Lys Lys Ser Leu Arg Xaa Leu Asn Tyr  
20 25 30

Ser Val Pro His Pro Met Leu Xaa His Pro Pro Pro Asp Thr Ala Gln  
35 40 45

40

(2) INFORMATION FOR SEQ ID NO: 155:

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 106 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

50

Met Gln Pro Leu Asn Phe Ser Ser Thr Glu Cys Ser Ser Phe Ser Pro  
1 5 10 15

55

Pro Thr Thr Val Ile Leu Ile Leu Cys Phe Glu Gly Leu Leu  
20 25 30

Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile  
35 40 45

60

Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg  
50 55 60

Trp Ala Lys Lys Thr Lys Trp Met Asn Met Lys Ala Val Phe Gly His  
65 70 75 80

5 Pro Phe Ser Leu Gly Trp Ala Ser Pro Phe Ala Thr Pro Asp Gln Gly  
85 90 95

Lys Ala Asp Pro Tyr Gln Tyr Val Val Xaa  
100 105

10

## (2) INFORMATION FOR SEQ ID NO: 156:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

20 Met Tyr Thr Asn His Phe Asn Leu Tyr Leu Lys Tyr Ile Leu Leu Ile  
1 5 10 15

25 Ile Leu Ile Leu Asn Met Thr Asn Ser Ser Ser Arg Tyr  
20 25

## (2) INFORMATION FOR SEQ ID NO: 157:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 53 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

Met Asn Glu Leu Leu Phe Phe Phe Phe Phe Phe Phe Thr Phe  
1 5 10 15

40 Cys Ile Glu Thr Asn Ser Phe Lys Gln Thr Tyr Tyr Tyr Phe Leu  
20 25 30

Gln Asn Ile Tyr Met Glu Met Leu Pro Pro Pro Val Asn Pro Pro Val  
35 40 45

45 Pro Pro Trp Gly Xaa  
50

50

## (2) INFORMATION FOR SEQ ID NO: 158:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 75 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

60 Met Tyr Ala Val Tyr Gln Gln Leu Ala Gln Leu Thr Leu Met Val Thr  
1 5 10 15

Leu Leu Ala Pro Ile Leu Pro Asp Glu Gln Ser Glu Val Phe Glu Ala  
20 25 30

5 Leu Ser Asn Leu Pro Lys Val Thr Trp Leu Gly Ser Asn Ser Pro Ser  
35 40 45

Ser Glu Met Pro Glu Pro Gly Arg Phe Val Ile Val His His Gln Leu  
50 55 60

10 Ser Ala Ala Ser His Ser Ser Ser Gln Leu Ala  
65 70 75

15

(2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:  
20 (A) LENGTH: 81 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

Met Trp Pro Pro Leu Leu Leu Leu Leu Leu Pro Ala Ala Pro  
25 1 5 10 15

Val Pro Thr Ala Lys Ala Ala Pro His Pro Asp Ala Asn Thr Gln Glu  
20 25 30

30 Gly Leu Gln Asn Leu Leu Gln Gly Val Gly Ala Gly Gly Asp Gly Glu  
35 40 45

Leu Arg Ala Asp Ser His Leu Ala Pro Gly Ser Gly Cys Ile Asp Gly  
50 55 60

35 Ala Val Val Ala Thr Arg Pro Glu Ser Arg Gly Gly Arg Pro Ala Val  
65 70 75 80

40 Pro

45

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:  
50 (A) LENGTH: 139 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu  
1 5 10 15

55 Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala  
20 25 30

Gln Glu Ala Gly Thr Ser Lys Pro Asn Glu Glu Ile Ser Gly Pro Ala  
35 40 45

60

295

	Glu Pro Ala Ser Pro Pro Glu Thr Thr Thr Ala Gln Glu Thr Ser
	50 55 60
5	Ala Ala Ala Val Gln Gly Thr Ala Lys Val Thr Ser Ser Arg Gln Glu
	65 70 75 80
	Leu Asn Pro Leu Lys Ser Ile Val Glu Lys Ser Ile Leu Leu Thr Glu
	85 90 95
10	Gln Ala Leu Ala Lys Ala Gly Lys Gly Met His Gly Gly Val Pro Gly
	100 105 110
	Gly Lys Gln Phe Ile Glu Asn Gly Ser Glu Phe Ala Gln Lys Leu Leu
	115 120 125
15	Lys Lys Phe Ser Leu Leu Lys Pro Trp Ala Xaa
	130 135

20

(2) INFORMATION FOR SEQ ID NO: 161:

	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 178 amino acids
25	(B) TYPE: amino acid
	(D) TOPOLOGY: linear
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:
30	Met Leu Gly Cys Gly Ile Pro Ala Leu Gly Leu Leu Leu Leu Gln
	1 5 10 15
	Gly Ser Ala Asp Gly Asn Gly Ile Gln Gly Phe Phe Tyr Pro Trp Ser
	20 25 30
35	Cys Glu Gly Asp Ile Trp Asp Arg Glu Ser Cys Gly Gly Gln Ala Ala
	35 40 45
	Ile Asp Ser Pro Asn Leu Cys Leu Arg Leu Arg Cys Cys Tyr Arg Asn
	50 55 60
40	Gly Val Cys Tyr His Gln Arg Pro Asp Glu Asn Val Arg Arg Lys His
	65 70 75 80
45	Met Trp Ala Leu Val Trp Thr Cys Ser Gly Leu Leu Leu Ser Cys
	85 90 95
	Ser Ile Cys Leu Phe Trp Trp Ala Lys Arg Arg Asp Val Leu His Met
	100 105 110
50	Pro Gly Phe Leu Ala Gly Pro Cys Asp Met Ser Lys Ser Val Ser Leu
	115 120 125
	Leu Ser Lys His Arg Gly Thr Lys Lys Thr Pro Ser Thr Gly Ser Val
	130 135 140
55	Pro Val Ala Leu Ser Lys Glu Ser Arg Asp Val Glu Gly Gly Thr Glu
	145 150 155 160
60	Gly Glu Gly Thr Glu Gly Glu Glu Thr Glu Gly Glu Glu Glu Glu
	165 170 175

Asp Xaa

5

(2) INFORMATION FOR SEQ ID NO: 162:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 72 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

15 Met Glu Ala Val Phe Thr Val Phe Phe Val Val Val Leu Phe Leu  
1 5 10 15

Lys Asn Thr Glu Gly Ala Lys Leu Phe Cys Thr Leu Tyr Pro Ala Ala  
20 25 30

20 Ser Ser Gly Gln Ser Gln Gly Pro Gly Leu Glu Lys Pro Asp Ser Gln  
35 40 45

25 Glu Cys Ile Ile Asp Pro Cys Ser Tyr Pro Ile Ala Leu Gly Ala Gly  
50 55 60

30 Thr Glu Pro Gly Cys Lys Ile Xaa  
65 70

30

(2) INFORMATION FOR SEQ ID NO: 163:

35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 67 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

40 Met Trp Phe Tyr Phe Leu Ser Val Phe Pro Leu Leu Pro Val Xaa  
1 5 10 15

Ala Pro Xaa Pro Pro Pro Ala Pro Thr Thr Leu Cys Leu Leu Phe  
20 25 30

45 Leu Gly Xaa Leu Tyr Asn Ser Thr Cys Ile His Cys Val His Thr Thr  
35 40 45

50 Ser Xaa Thr Gln Asn Pro Thr Ala Asn Thr Leu Lys Lys Lys Lys  
50 55 60

55 Asn Trp Gly  
65

55

(2) INFORMATION FOR SEQ ID NO: 164:

60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 155 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

5 Met Gly Phe Gly Ala Thr Leu Ala Val Gly Leu Thr Ile Phe Val Leu  
   1               5                           10                   15

10 Ser Val Val Thr Ile Ile Ile Cys Phe Thr Cys Ser Cys Cys Cys Leu  
     20                                   25                           30

15 Tyr Lys Thr Cys Arg Arg Pro Arg Pro Val Val Thr Thr Thr Thr Ser  
     35                                   40                           45

20 Thr Thr Val Val His Ala Pro Tyr Pro Gln Pro Pro Ser Val Pro Pro  
     50                                   55                           60

25 Ser Tyr Pro Gly Pro Ser Tyr Gln Gly Tyr His Thr Met Pro Pro Gln  
     65                                   70                           75                           80

30 Pro Gly Met Pro Ala Ala Pro Tyr Pro Met Gln Tyr Pro Pro Pro Tyr  
     85                                   90                           95

35 Pro Ala Gln Pro Met Gly Pro Pro Ala Tyr His Glu Thr Leu Ala Gly  
     100                                   105                           110

40 Glu Gln Pro Arg Pro Thr Pro Pro Ala Ser Leu Leu Thr Thr Arg Pro  
     115                                   120                           125

45 Thr Trp Met Pro Arg Arg Pro Ser Glu His Ser Leu Ala Ser Leu  
     130                                   135                           140

50 Ala Ala Thr Trp Leu Cys Cys Val Cys Ala Xaa  
     145                                   150                           155

55

(2) INFORMATION FOR SEQ ID NO: 165:

40 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 104 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

50 Met Ile Ile Leu Val Phe Ile Ala Phe Phe Ile Pro Leu Gln Lys Thr  
     1               5                           10                   15

55 Ile Gly Lys Ile Ala Thr Cys Leu Glu Leu Arg Ser Ala Ala Leu Gln  
     20                                   25                           30

60 Ser Thr Gln Ser Gln Glu Glu Phe Lys Leu Glu Asp Leu Lys Lys Leu  
     35                                   40                           45

65 Glu Pro Ile Leu Lys Asn Ile Leu Thr Tyr Asn Lys Glu Phe Pro Phe  
     50                                   55                           60

70 Asp Val Gln Pro Val Pro Leu Arg Arg Ile Leu Ala Pro Gly Glu Glu  
     65                                   70                           75                           80

75 Glu Asn Leu Glu Phe Glu Asp Glu Glu Gly Gly Ala Gly Ala

85

90

95

Gly Leu Leu Ile Leu Ser Cys Xaa  
 100

5

## (2) INFORMATION FOR SEQ ID NO: 166:

10                   (i) SEQUENCE CHARACTERISTICS:  
                       (A) LENGTH: 81 amino acids  
                       (B) TYPE: amino acid  
                       (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:  
 15                   Met Ala Gly Thr Met Val Ile Val Val Val Val Val Gly Glu Val  
                       1                   5                   10                   15  
 20                   Val Val Glu Ala Glu Val Val Val Gln Ala Arg Glu Glu Ala Gly Glu  
                       20                   25                   30  
 25                   Glu Glu Gly Ala Arg Ile Ile Thr Lys Gly Val Asn Leu Asn Ser Ile  
                       35                   40                   45  
 30                   Ser Ser Met Glu Val Ile Ser Ile Ile Ile Leu Asp Leu Asp Arg Glu  
                       50                   55                   60  
 35                   Asp Ile Thr Leu Val Glu Ala Thr Glu Pro Tyr Ile Leu Leu Glu Leu  
                       65                   70                   75                   80  
 40                   Lys  
 45                   

35

## (2) INFORMATION FOR SEQ ID NO: 167:

40                   (i) SEQUENCE CHARACTERISTICS:  
                       (A) LENGTH: 93 amino acids  
                       (B) TYPE: amino acid  
                       (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:  
 45                   Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile  
                       1                   5                   10                   15  
 50                   Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys  
                       20                   25                   30  
 55                   Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val  
                       35                   40                   45  
 60                   Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu  
                       50                   55                   60  
 65                   Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser  
                       70                   75                   80  
 70                   Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr  
                       85                   90  
 75

## (2) INFORMATION FOR SEQ ID NO: 168:

5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

Met Gly Trp Ser Ala Gly Leu Leu Phe Leu Leu Ile Leu Tyr Leu Pro  
1 5 10 15

15

Val Pro Gly Trp Met Glu Arg Glu Asp Gly Glu Thr Gly His Leu Ser  
20 25 30

Pro Gln Ala Pro Gly Arg Glu Tyr Arg Gly Phe Tyr Ser Val Pro Pro  
35 40 45

20

Asp Tyr Val Trp Leu Arg Asp Ser Pro Xaa  
50 55

25

## (2) INFORMATION FOR SEQ ID NO: 169:

30

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 232 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

35

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser  
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Cys Gln Thr  
20 25 30

40

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro  
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys  
50 55 60

45

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro  
65 70 75 80

50

Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg  
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu  
100 105 110

55

Gly Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile  
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp  
130 135 140

60

300

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala  
145 150 155 160

5 Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Thr  
165 170 175

Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu  
180 185 190

10 Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr  
195 200 205

Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu  
210 215 220

15 Gly Phe Ile Leu Ile Pro Cys Xaa  
225 230

20

(2) INFORMATION FOR SEQ ID NO: 170:

- (i) SEQUENCE CHARACTERISTICS:  
25 (A) LENGTH: 72 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

30 Met Ser Ala Ile Phe Asn Phe Gln Ser Leu Leu Thr Val Ile Leu Leu  
1 5 10 15

Leu Ile Cys Thr Cys Ala Tyr Ile Arg Ser Leu Ala Pro Ser Leu Leu  
20 25 30

35 Asp Arg Asn Lys Thr Gly Leu Leu Gly Ile Phe Trp Lys Cys Ala Arg  
35 40 45

Ile Gly Glu Arg Lys Ser Pro Tyr Val Ala Val Cys Cys Ile Val Met  
50 55 60

40 Ala Phe Ser Ile Leu Phe Ile Gln  
65 70

45

(2) INFORMATION FOR SEQ ID NO: 171:

- (i) SEQUENCE CHARACTERISTICS:  
50 (A) LENGTH: 65 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

55 Met Gly Thr Phe Ser Leu Ser Leu Phe Gly Leu Met Gly Val Ala Phe  
1 5 10 15

Gly Met Asn Leu Glu Ser Ser Leu Glu Glu Asp His Arg Ile Phe Trp  
20 25 30

60 Leu Ile Thr Gly Ile Met Phe Met Gly Ser Gly Leu Ile Trp Arg Arg

35

40

45

Leu Leu Ser Phe Leu Gly Arg Gln Leu Glu Ala Pro Leu Pro Pro Met  
 50 55 60

5

Val  
 65

10

(2) INFORMATION FOR SEQ ID NO: 172:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

Met Tyr Lys Gly Lys Leu Val Ile Val Leu Ile Leu Leu Leu Pro  
 20 1 5 10 15

Ser His Phe Met Phe Leu Thr Gln Cys Lys Glu Ile Lys His Asn Leu  
 20 25 30

25 Lys Lys Asn Met Ser Leu Leu Leu Phe Thr Ile Lys Ser Trp Leu Tyr  
 35 40 45

Ser Ala Ser Leu Gly Ile Leu Tyr Asn Trp Gln His Leu Thr Ala Gln  
 50 55 60

30 Val Asp Gln Cys Thr Ser Leu Ile Leu Ile His  
 65 70 75

35

(2) INFORMATION FOR SEQ ID NO: 173:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 334 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

40 Met Val Gly His Glu Met Ala Ser Xaa Ser Ser Asn Thr Ser Leu Pro  
 45 1 5 10 15

Phe Ser Asn Met Gly Asn Pro Met Asn Thr Thr Gln Leu Gly Lys Ser  
 20 25 30

50 Leu Phe Gln Trp Gln Val Glu Gln Glu Ser Lys Leu Ala Asn Ile  
 35 40 45

Ser Gln Asp Gln Phe Leu Ser Lys Asp Ala Asp Gly Asp Thr Phe Leu  
 50 55 60

55 His Ile Ala Val Ala Gln Gly Arg Arg Ala Leu Ser Tyr Val Leu Ala  
 65 70 75 80

60 Arg Lys Met Asn Ala Leu His Met Leu Asp Ile Lys Glu His Asn Gly  
 85 90 95

Gln Ser Ala Phe Gln Val Ala Val Ala Asn Gln His Leu Ile Val  
 100 105 110

5 Gln Asp Leu Val Asn Ile Gly Ala Gln Val Asn Thr Thr Asp Cys Trp  
 115 120 125

Gly Arg Thr Pro Leu His Val Cys Ala Glu Lys Gly His Ser Gln Val  
 130 135 140

10 Leu Gln Ala Ile Gln Lys Gly Ala Val Gly Ser Asn Gln Phe Val Asp  
 145 150 155 160

15 Leu Glu Ala Thr Asn Tyr Asp Gly Leu Thr Pro Leu His Cys Ala Val  
 165 170 175

Ile Ala His Asn Ala Val Val His Glu Leu Gln Arg Asn Gln Gln Pro  
 180 185 190

20 His Ser Pro Glu Val Gln Glu Leu Leu Leu Lys Asn Lys Ser Leu Val  
 195 200 205

Asp Thr Ile Lys Cys Leu Ile Gln Met Gly Ala Ala Val Glu Ala Lys  
 210 215 220

25 Asp Arg Lys Ser Gly Arg Thr Ala Leu His Leu Ala Ala Glu Glu Ala  
 225 230 235 240

30 Asn Leu Glu Leu Ile Arg Leu Phe Leu Glu Leu Pro Ser Cys Leu Ser  
 245 250 255

Phe Val Asn Ala Lys Ala Tyr Asn Gly Asn Thr Ala Leu His Val Ala  
 260 265 270

35 Ala Ser Leu Gln Tyr Arg Leu Thr Gln Leu Asp Ala Val Arg Leu Leu  
 275 280 285

Met Arg Lys Gly Ala Asp Pro Ser Thr Arg Asn Leu Glu Asn Glu Gln  
 290 295 300

40 Pro Val His Leu Val Pro Asp Gly Pro Val Gly Glu Gln Ile Arg Arg  
 305 310 315 320

Ile Leu Lys Gly Lys Ser Ile Gln Gln Arg Ala Pro Pro Tyr  
 45 325 330

(2) INFORMATION FOR SEQ ID NO: 174:  
 50

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 196 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

Met Asp Ala Arg Trp Trp Ala Val Val Val Leu Ala Ala Phe Pro Ser  
 1 5 10 15

60 Leu Gly Ala Gly Gly Glu Thr Pro Glu Ala Pro Pro Glu Ser Trp Thr

	20	25	30
	Gln Leu Trp Phe Phe Arg Phe Val Val Asn Ala Ala Gly Tyr Ala Ser		
	35	40	45
5	Phe Met Val Pro Gly Tyr Leu Leu Val Gln Tyr Phe Arg Arg Lys Asn		
	50	55	60
10	Tyr Leu Glu Thr Gly Arg Gly Leu Cys Phe Pro Leu Val Lys Ala Cys		
	65	70	75
	Val Phe Gly Asn Glu Pro Lys Ala Ser Asp Glu Val Pro Leu Ala Pro		
	85	90	95
15	Arg Thr Glu Ala Ala Glu Thr Thr Pro Met Trp Gln Ala Leu Lys Leu		
	100	105	110
	Leu Phe Cys Ala Thr Gly Leu Gln Val Ser Tyr Leu Thr Trp Gly Val		
	115	120	125
20	Leu Gln Glu Arg Val Met Thr Arg Ser Tyr Gly Ala Thr Ala Thr Ser		
	130	135	140
25	Pro Gly Glu Arg Phe Thr Asp Ser Gln Phe Leu Val Leu Met Asn Arg		
	145	150	155
	Val Leu Ala Leu Ile Val Ala Gly Leu Ser Cys Val Leu Cys Lys Gln		
	165	170	175
30	Pro Arg His Gly Ala Pro Met Tyr Arg Tyr Ser Phe Cys Gln Pro Val		
	180	185	190
	Gln Cys Ala Xaa		
	195		
35			

## (2) INFORMATION FOR SEQ ID NO: 175:

40	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 265 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:		
45	Met Ser Asp Leu Leu Leu Gly Leu Ile Gly Gly Leu Thr Leu Leu		
	1	5	10
	15		
50	Leu Leu Leu Thr Leu Leu Ala Phe Ala Gly Tyr Ser Gly Leu Leu Ala		
	20	25	30
	30		
	Gly Val Glu Val Ser Ala Gly Ser Pro Pro Ile Arg Asn Val Thr Val		
	35	40	45
55	Ala Tyr Lys Phe His Met Gly Leu Tyr Gly Glu Thr Gly Arg Leu Phe		
	50	55	60
	Thr Glu Ser Cys Ser Ile Ser Pro Lys Leu Arg Ser Ile Ala Val Tyr		
	65	70	75
	80		
60			

304

	Tyr Asp Asn Pro His Met Val Pro Pro Asp Lys Cys Arg Cys Ala Val			
	85	90	95	
	Gly Ser Ile Leu Ser Glu Gly Glu Ser Pro Ser Pro Glu Leu Ile			
5	100	105	110	
	Asp Leu Tyr Gln Lys Phe Gly Phe Lys Val Phe Ser Phe Pro Glu Pro			
	115	120	125	
10	Ser His Val Val Thr Ala Thr Phe Pro Leu Thr Pro Pro Phe Cys Pro			
	130	135	140	
	Ile Trp Leu Gly Tyr Pro Pro Cys Pro Ser Cys Leu Gly His Leu His			
	145	150	155	160
15	Gln Gly Ala Glu Ala Val Cys Leu Ser Ser Ala Gly Asp Leu Pro Gly			
	165	170	175	
20	Arg Pro Glu Ser Ile Ser Cys Ala His Trp His Gly Gln Gly Asp Phe			
	180	185	190	
	Tyr Val Pro Glu Met Lys Glu Thr Glu Trp Lys Trp Arg Gly Leu Val			
	195	200	205	
25	Glu Ala Ile Asp Thr Gln Val Asp Gly Thr Gly Ala Asp Thr Met Ser			
	210	215	220	
	Asp Thr Ser Ser Val Ser Leu Glu Val Ser Pro Gly Ser Arg Glu Thr			
	225	230	235	240
30	Ser Ala Ala Thr Leu Ser Pro Gly Ala Ser Ser Arg Gly Trp Asp Asp			
	245	250	255	
	Gly Asp Thr Arg Ser Glu His Ser Xaa			
35	260	265		

## (2) INFORMATION FOR SEQ ID NO: 176:

40

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

	Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln			
	1	5	10	15
50	Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp			
	20	25	30	
	Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val			
	35	40	45	
55	Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro			
	50	55	60	
60	Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Trp Thr			
	65	70	75	80

305

Phe Leu Ser Arg Gly Arg Glu Ala Glu Val Leu Val Ala Arg Gly Val  
 85 90 95

5 Arg Val Lys Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala  
 100 105 110

Tyr Pro Ala Ser Leu Thr Asp Val Ser Pro Gly Ala Glu Arg Ala Ala  
 115 120 125

10 Pro Gln Arg Leu Arg Tyr Leu Ser Leu Xaa  
 130 135

15 (2) INFORMATION FOR SEQ ID NO: 177:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 179 amino acids

20 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:

25 Met Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp  
 1 5 10 15

Leu Cys Cys Ala Thr Pro Ala His Ala Leu Gln Cys Arg Asp Gly Tyr  
 20 25 30

30 Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn Gly Thr  
 35 40 45

Gly Tyr Cys Lys Gly Pro Glu Gly Phe Leu Gly Glu Tyr Cys Gln His  
 50 55 60

35 Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr Cys Val  
 65 70 75 80

40 Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser Gly Phe  
 85 90 95

Thr Gly Glu Asp Cys Gln Tyr Ser Thr Ser His Pro Cys Phe Val Ser  
 100 105 110

45 Arg Pro Cys Leu Asn Gly Gly Thr Cys His Met Leu Ser Arg Asp Thr  
 115 120 125

Tyr Glu Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys Gln Trp  
 130 135 140

50 Thr Asp Ala Cys Leu Ser His Pro Cys Ala Asn Gly Ser Thr Cys Thr  
 145 150 155 160

55 Thr Val Ala Asn His Phe Leu Gln Met Pro His Arg Leu His Arg Ala  
 165 170 175

Glu Val Xaa

60

(2) INFORMATION FOR SEQ ID NO: 178:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 155 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

10 Met Thr Arg Gly Gly Pro Gly Gly Arg Pro Gly Leu Pro Gln Pro Pro  
1 5 10 15

Pro Leu Leu Leu Leu Leu Leu Pro Leu Leu Leu Val Thr Ala Glu  
20 25 30

15 Pro Pro Lys Pro Ala Gly Val Tyr Tyr Ala Thr Ala Tyr Trp Met Pro  
35 40 45

20 Ala Glu Lys Thr Val Gln Val Lys Asn Val Met Asp Lys Asn Gly Asp  
50 55 60

Ala Tyr Gly Phe Tyr Asn Asn Ser Val Lys Thr Thr Gly Trp Gly Ile  
65 70 75 80

25 Leu Glu Ile Arg Ala Gly Tyr Gly Ser Gln Thr Leu Ser Asn Glu Ile  
85 90 95

Ile Met Phe Val Ala Gly Phe Leu Glu Gly Tyr Leu Ile Ala Pro His  
100 105 110

35            Ser Ile Met Asp Lys Val Gln Asp Phe Met Glu Lys Gln Asp Lys Val  
               130                            135                                    140

Asp Pro Glu Lys Tyr Gln Arg Ile Gln Asp Xaa  
 145 150 155

40

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

50 Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His  
1 5 10 15

Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp  
20 25 30

55 Phe Ser Tyr Lys Arg Xaa Asn Cys Lys Pro Ile Pro Val Asn Leu Gln  
35 40 45

60 Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu  
50 55 60

	Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile	65	70	75	80
5	Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys	85	90	95	
	Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln	100	105	110	
10	Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val	115	120	125	
	Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg	130	135	140	
15	Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His	145	150	155	160
20	Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys	165	170	175	
	Asn Lys Asn Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn	180	185	190	
25	Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg	195	200	205	
	Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu	210	215	220	
30	Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys	225	230	235	240
	Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro	245	250	255	
35	Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser	260	265	270	
	Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg	275	280	285	
40	Ser Ile Arg Lys Leu Gln Cys	290	295		
45					

(2) INFORMATION FOR SEQ ID NO: 180:

50 (i) SEQUENCE CHARACTERISTICS:  
      (A) LENGTH: 256 amino acids  
      (B) TYPE: amino acid  
      (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

Met Arg Pro Ala Ala Leu Arg Gly Ala Leu Leu Gly Cys Leu Cys Leu  
 1                   5                   10                   15  
 Ala Leu Leu Cys Leu Gly Gly Ala Asp Lys Arg Leu Arg Asp Asn His

308

	20	25	30
	Glu Trp Lys Lys Leu Ile Met Val Gln His Trp Pro Glu Thr Val Cys		
	35	40	45
5	Glu Lys Ile Gln Asn Asp Cys Arg Asp Pro Pro Asp Tyr Trp Thr Ile		
	50	55	60
	His Gly Leu Trp Pro Asp Lys Ser Glu Gly Cys Asn Arg Ser Trp Pro		
10	65	70	75
	Phe Asn Leu Glu Glu Ile Lys Asp Leu Leu Pro Glu Met Arg Ala Tyr		
	85	90	95
15	Trp Pro Asp Val Ile His Ser Phe Pro Asn Arg Ser Arg Phe Trp Lys		
	100	105	110
	His Glu Trp Glu Lys His Gly Thr Cys Ala Ala Gln Val Asp Ala Leu		
	115	120	125
20	Asn Ser Gln Lys Lys Tyr Phe Gly Arg Ser Leu Glu Leu Tyr Arg Glu		
	130	135	140
	Leu Asp Leu Asn Ser Val Leu Leu Lys Leu Gly Ile Lys Pro Ser Ile		
25	145	150	155
	Asn Tyr Tyr Gln Val Ala Asp Phe Lys Asp Ala Leu Ala Arg Val Tyr		
	165	170	175
30	Gly Val Ile Pro Lys Ile Gln Cys Leu Pro Pro Ser Gln Asp Glu Glu		
	180	185	190
	Val Gln Thr Ile Gly Gln Ile Glu Leu Cys Leu Thr Lys Gln Asp Gln		
	195	200	205
35	Gln Leu Gln Asn Cys Thr Glu Pro Gly Glu Gln Pro Ser Pro Lys Gln		
	210	215	220
	Glu Val Trp Leu Ala Asn Gly Ala Ala Glu Ser Arg Gly Leu Arg Val		
40	225	230	235
	Cys Glu Asp Gly Pro Val Phe Tyr Pro Pro Pro Lys Lys Thr Lys His		
	245	250	255
45			

50 (2) INFORMATION FOR SEQ ID NO: 181:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 324 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

Met Ala Pro Leu Leu Gln Leu Ala Val Leu Gly Ala Ala Leu Ala		
1	5	10
		15

60

Ala Ala Ala Leu Val Leu Ile Ser Ile Val Ala Phe Thr Thr Ala Thr  
 20 25 30

Lys Met Pro Ala Leu His Arg His Glu Glu Glu Lys Phe Phe Leu Asn  
 5 35 40 45

Ala Lys Gly Gln Lys Glu Thr Leu Pro Ser Ile Trp Asp Ser Pro Thr  
 50 55 60

10 Lys Gln Leu Ser Val Val Pro Ser Tyr Asn Glu Glu Lys Arg Leu  
 65 70 75 80

Pro Val Met Met Asp Glu Ala Leu Ser Tyr Leu Glu Lys Arg Gln Lys  
 85 90 95

15 Arg Asp Pro Ala Phe Thr Tyr Glu Val Ile Val Val Asp Asp Gly Ser  
 100 105 110

20 Lys Asp Gln Thr Ser Lys Val Ala Phe Lys Tyr Cys Gln Lys Tyr Gly  
 115 120 125

Ser Asp Lys Val Arg Val Ile Thr Leu Val Lys Asn Arg Gly Lys Gly  
 130 135 140

25 Gly Ala Ile Arg Met Gly Ile Phe Ser Ser Arg Gly Glu Lys Ile Leu  
 145 150 155 160

Met Ala Asp Ala Asp Gly Ala Thr Lys Phe Pro Asp Val Glu Lys Leu  
 165 170 175

30 Glu Lys Gly Leu Asn Asp Leu Gln Pro Trp Pro Asn Gln Met Ala Ile  
 180 185 190

Ala Cys Gly Ser Arg Ala His Leu Glu Lys Glu Ser Ile Ala Gln Arg  
 35 195 200 205

Ser Tyr Phe Arg Thr Leu Leu Met Tyr Gly Phe His Phe Leu Val Trp  
 210 215 220

40 Phe Leu Cys Val Lys Gly Ile Arg Asp Thr Gln Cys Gly Phe Lys Leu  
 225 230 235 240

Phe Thr Arg Glu Ala Ala Ser Arg Thr Phe Ser Ser Leu His Val Glu  
 245 250 255

45 Arg Trp Ala Phe Asp Val Glu Leu Leu Tyr Ile Ala Gln Phe Phe Lys  
 260 265 270

Ile Pro Ile Ala Glu Ile Ala Val Asn Trp Thr Glu Ile Glu Gly Ser  
 50 275 280 285

Lys Leu Val Pro Phe Trp Ser Trp Leu Gln Met Gly Lys Asp Leu Leu  
 290 295 300

55 Phe Ile Arg Leu Arg Tyr Leu Thr Gly Ala Trp Arg Leu Glu Gln Thr  
 305 310 315 320

Arg Lys Met Asn

## (2) INFORMATION FOR SEQ ID NO: 182:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

10 Met Asp Ile Cys Phe Phe His Tyr Val Leu Leu Phe Phe Leu Val Arg  
1 5 10 15

Cys Ala Leu Val Val Leu Ile Leu Leu Cys Gln Gly Trp Gly Asn Gly  
15 20 25 30

Gly Gly Cys Val Gly Arg Val Leu Ile Ile Val Phe Ser Ser Val  
35 40 45

20

## (2) INFORMATION FOR SEQ ID NO: 183:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

30 Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr  
1 5 10 15

Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp  
20 25 30

Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe  
35 40 45

Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe  
40 50 55 60

Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Glu Gly Thr Phe  
65 70 75 80

45 Arg Ser Ser Ile Arg Arg Leu Ser Thr Arg Arg Arg Xaa  
85 90

50 (2) INFORMATION FOR SEQ ID NO: 184:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 168 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

Met Xaa Thr Lys Glu Phe Gly Xaa Gly Arg Ala Val Gln Gln Val Leu  
1 5 10 15

60

Asn Ile Glu Cys Leu Arg Asp Phe Leu Thr Pro Pro Leu Leu Ser Val  
 20 25 30  
 Arg Phe Arg Tyr Val Gly Ala Pro Gln Ala Leu Thr Leu Lys Leu Pro  
 5 35 40 45  
 Val Thr Xaa Asn Lys Phe Phe Gln Pro Thr Glu Met Ala Ala Gln Asp  
 50 55 60  
 10 Phe Phe Gln Arg Trp Lys Gln Leu Ser Leu Pro Gln Gln Glu Ala Gln  
 65 70 75 80  
 Lys Ile Phe Lys Ala Asn His Pro Met Asp Ala Glu Val Thr Lys Ala  
 85 90 95  
 15 Lys Leu Leu Gly Phe Gly Ser Ala Leu Leu Asp Asn Val Asp Pro Asn  
 100 105 110  
 20 Pro Glu Asn Phe Val Gly Ala Gly Ile Ile Gln Thr Lys Ala Leu Gln  
 115 120 125  
 Val Gly Cys Leu Leu Arg Leu Glu Pro Asn Ala Gln Ala Gln Met Tyr  
 130 135 140  
 25 Arg Leu Thr Leu Arg Thr Ser Lys Glu Pro Val Ser Arg His Leu Cys  
 145 150 155 160  
 Glu Leu Leu Ala Gln Gln Phe Xaa  
 165  
 30

## (2) INFORMATION FOR SEQ ID NO: 185:

35 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 43 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 185:  
 40 Met Phe Tyr Val Leu Ser Val Ser Pro Leu Leu Xaa Phe Leu Ala Cys  
 1 5 10 15  
 Gly Leu Cys Leu Cys Val Asn Trp Lys Ile Ala Ile Ser Gln Leu Ser  
 45 20 25 30  
 Leu Ser Phe Lys Asn Glu Leu Glu Lys Pro Xaa  
 35 40  
 50

## (2) INFORMATION FOR SEQ ID NO: 186:

55 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 59 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:  
 60 Met Lys Leu Phe Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly

312

1	5	10	15
His Ile Leu Ala Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu			
	20	25	30
5	Gly Pro Gly Trp Val Pro Ser Ala Leu Xaa Arg Leu His Pro Gly His		
	35	40	45
Leu Ser Gly Ser Val Leu Val Ser Ala Ala Xaa			
10	50	55	

(2) INFORMATION FOR SEQ ID NO: 187:

15	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 189 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:		
Met Asp Val Asn Ile Ala Pro Leu Arg Ala Trp Asp Asp Phe Phe Pro			
	1	5	10
	15		
25	Gly Ser Asp Arg Phe Ala Arg Pro Asp Phe Arg Asp Ile Ser Lys Trp		
	20	25	30
Asn Asn Arg Val Val Ser Asn Leu Leu Tyr Tyr Gln Thr Asn Tyr Leu			
	35	40	45
30	Val Val Ala Ala Met Met Ile Ser Ile Val Gly Phe Leu Ser Pro Phe		
	50	55	60
Asn Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe			
35	65	70	75
	80		
Val Trp Ala Ala His Asn Lys Asp Val Leu Arg Arg Met Lys Lys Arg			
	85	90	95
40	Tyr Pro Thr Thr Phe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu		
	100	105	110
Ile Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe			
	115	120	125
45	Pro Leu Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu		
	130	135	140
Lys Asn Lys Leu Glu Asn Lys Met Glu Gly Ile Gly Leu Lys Arg Thr			
50	145	150	155
	160		
Pro Met Gly Ile Val Leu Asp Ala Leu Glu Gln Gln Glu Gly Ile			
	165	170	175
55	Asn Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu Xaa		
	180	185	

60 (2) INFORMATION FOR SEQ ID NO: 188:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 146 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

Met Phe Leu Thr Arg Ile Leu Cys Pro Thr Tyr Ile Ala Leu Thr Phe  
 1               5                           10                           15

10

Leu Val Tyr Ile Val Ala Leu Val Ser Gly Gln Leu Cys Met Glu Ile  
 20   25                                   30

15

Ala Arg Gly Asn Ile Phe Phe Leu Asn Glu Leu Val Thr Thr Phe Cys  
 35   40                                   45

Cys Ser Cys Leu Leu Ser Val Pro Tyr Leu His Pro Gly Phe Phe  
 50   55                                   60

20

Tyr Ser Ser Leu Cys Lys Cys Cys Phe Val Leu Val Val Leu Ser Arg  
 65   70                                   75                                   80

Ile Gly Ser Val Asn Glu Thr Trp Ser Cys Asn Phe Ser Ile Cys Ser  
 85   90                                   95

25

Tyr Leu Ile Phe Gly Ser Pro Ile Phe Thr Ala Val Ile Pro Lys Arg  
 100   105                                   110

30

Cys Ala Leu Glu Asp Ile Gln Asn Asn Pro Ile Gly Cys Leu Leu Arg  
 115   120                                   125

Cys Thr Pro Ala Trp Glu Thr Glu Gly Asp Ser Ile Ser Lys Lys Ile  
 130   135                                   140

35

Lys Lys  
 145

40

## (2) INFORMATION FOR SEQ ID NO: 189:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

Met Gly Ser Arg Ala Glu Leu Cys Thr Leu Leu Gly Gly Phe Ser Phe  
 1               5                           10                           15

50

Leu Leu Leu Leu Ile Pro Gly Glu Gly Ala Lys Gly Ser Leu Arg  
 20   25                                   30

55

Glu Ser Gln Gly Val Cys Ser Lys Gln Thr Leu Val Val Pro Leu His  
 35   40                                   45

Tyr Asn Glu Ser Tyr Ser Gln Pro Val Tyr Lys Pro Tyr Leu Thr Leu  
 50   55                                   60

60

Cys Ala Gly Ser Ala Ser Ala Ala Leu Thr Gly Pro Cys Thr Ala Leu

314

65 70 75 80

Cys Gly Gly Arg

5

## (2) INFORMATION FOR SEQ ID NO: 190:

## 10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

15 Met Met Gly Val Leu Gln Leu Leu His Ile Phe Trp Ala Tyr Leu Ile  
1 5 10 15

Leu Arg Met Ala His Lys Phe Ile Thr Gly Lys Leu Val Glu Asp Glu  
20 25 30

Arg Ser Thr Gly Lys Lys Gln Arg Ala Gln Arg Gly Arg Arg Leu Gln  
35 40 45

25 Leu Gly Glu Glu Gln Arg Ala Gly Pro Xaa  
50 55

## 30 (2) INFORMATION FOR SEQ ID NO: 191:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

35 Met Arg Arg Leu Val His Asp Leu Leu Pro Pro Glu Val Cys Ser Leu  
1 5 10 15

40 Leu Asn Pro Ala Ala Ile Tyr Ala Asn Asn Glu Ile Ser Leu Arg Asp  
20 25 30

45 Val Glu Val Tyr Gly Phe Asp Tyr Asp Tyr Thr Leu Ala Gln Tyr Ala  
35 40 45

Asp Ala Leu His Pro Glu Ile Phe Ser Thr Ala Arg Asp Ile Leu Ile  
50 55 60

50 Glu His Tyr Lys Tyr Pro Glu Gly Ile Arg Lys Tyr Asp Tyr Asn Pro  
65 70 75 80

Ser Phe Ala Ile Arg Gly Leu His Tyr Asp Ile Gln Lys Ser Leu Leu  
85 90 95

55 Met Lys Ile Asp Ala Phe His Tyr Val Gln Leu Gly Thr Ala Tyr Arg  
100 105 110

60 Gly Leu Gln Pro Val Pro Asp Glu Glu Val Ile Glu Leu Tyr Gly Gly  
115 120 125

Thr Gln His Ile Pro Leu Tyr Gln Met Ser Gly Phe Tyr Gly Lys Gly  
 130 135 140  
 5 Pro Ser Ile Lys Gln Phe Met Asp Ile Phe Ser Leu Pro Glu Met Ala  
 145 150 155 160  
 Leu Leu Ser Cys Val Val Asp Tyr Phe Leu Gly His Ser Leu Glu Phe  
 165 170 175  
 10 Asp Gln Ala His Leu Tyr Lys Asp Val Thr Asp Ala Ile Arg Asp Val  
 180 185 190  
 His Val Lys Gly Leu Met Tyr Gln Trp Ile Glu Gln Asp Met Glu Lys  
 15 195 200 205  
 Tyr Ile Leu Arg Gly Asp Glu Thr Phe Ala Val Leu Ser Arg Leu Val  
 210 215 220  
 20 Ala His Gly Lys Gln Leu Phe Leu Ile Thr Asn Ser Pro Phe Ser Phe  
 225 230 235 240  
 Val Asp Lys Gly Met Arg His Met Val Gly Pro Asp Trp Arg His Ser  
 245 250 255  
 25 Ser Met Trp Ser Leu Ser Arg Gln Thr Ser Pro Ala Ser Ser Leu Thr  
 260 265 270  
 Gly Ala Ser Phe Xaa Glu Asn Ser Met Arg Arg Ala His Phe Ser Gly  
 30 275 280 285  
 Thr Gly Ser Pro Ala Trp Lys Arg Ala Arg Ser Ile Gly Arg Glu Thr  
 290 295 300  
 35 Cys Leu Thr Ser Tyr Ala Xaa  
 305 310

40 (2) INFORMATION FOR SEQ ID NO: 192:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 318 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

Met Asn Trp Glu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu  
 1 5 10 15  
 50 Leu Leu Leu Val Gln Leu Leu Arg Phe Leu Arg Ala Asp Gly Asp Leu  
 20 25 30  
 Thr Leu Leu Trp Ala Glu Trp Gln Gly Arg Arg Pro Glu Trp Glu Leu  
 55 35 40 45  
 Thr Asp Met Val Val Trp Val Thr Gly Ala Ser Ser Gly Ile Gly Glu  
 50 55 60  
 60 Glu Leu Ala Tyr Gln Leu Ser Lys Leu Gly Val Ser Leu Val Leu Ser

316

	70	75	80
5	Ala Arg Arg Val His Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Glu 85	90	95
	Asn Gly Asn Leu Lys Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu 100	105	110
10	Thr Asp Thr Gly Ser His Glu Ala Ala Thr Lys Ala Val Leu Gln Glu 115	120	125
	Phe Gly Arg Ile Asp Ile Leu Val Asn Asn Gly Gly Met Ser Gln Arg 130	135	140
15	Ser Leu Cys Met Asp Thr Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu 145	150	155
	Leu Asn Tyr Leu Gly Thr Val Ser Leu Thr Lys Cys Val Leu Pro His 165	170	175
20	Met Ile Glu Arg Lys Gln Gly Lys Ile Val Thr Val Asn Ser Ile Leu 180	185	190
	Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala Ser Lys His 195	200	205
25	Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr 210	215	220
	Pro Gly Ile Ile Val Ser Asn Ile Cys Pro Gly Pro Val Gln Ser Asn 225	230	235
30	Ile Val Glu Asn Ser Leu Ala Gly Glu Val Thr Lys Thr Ile Gly Asn 245	250	255
	Asn Gly Asp Gln Ser His Lys Met Thr Thr Ser Arg Cys Val Arg Leu 260	265	270
35	Met Leu Ile Ser Met Ala Asn Asp Leu Lys Glu Val Trp Ile Ser Glu 275	280	285
	Gln Pro Phe Leu Phe Ser Asn Ile Phe Val Ala Ile His Ala Asn Leu 290	295	300
40	Gly Leu Val Asp Asn Gln Gln Asp Gly Glu Glu Lys Asp Xaa 305	310	315

50 (2) INFORMATION FOR SEQ ID NO: 193:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:

Met Trp Pro Ser Phe Pro Gln Val Arg Val Gly Ser Phe Leu Phe Gly  
1 5 10 15

60

Ile Leu Phe Phe Ser Phe Gly Ser Ser Ser Leu Pro Pro Gly Leu Pro  
 20 25 30

5 Pro Pro Ala Ser Leu Leu Cys Cys Ala Val Gln Trp Gly Ala Arg Ala  
 35 40 45

Leu Phe Leu Pro Ala  
 50

10

(2) INFORMATION FOR SEQ ID NO: 194:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 42 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

20 Met Leu Val Thr Cys Ser Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala  
 1 5 10 15

Ile Leu Ala Gln Leu Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu  
 20 25 30

25 Thr Ile Trp Tyr Leu Lys Tyr His Trp Pro  
 35 40

30

(2) INFORMATION FOR SEQ ID NO: 195:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 102 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

40 Met Glu Gly Thr Glu Met Gly Ala Arg Pro Gly Gly His Pro Gln Lys  
 1 5 10 15

Trp Ser Phe Leu Trp Ser Leu Ala Leu Trp Leu Pro Leu Ala Leu Ser  
 20 25 30

45 Val Ser Leu Phe Leu Gly Leu Ser Leu Ser Pro Pro Gln Pro Gly Leu  
 35 40 45

50 Ser Leu Trp Cys Thr Leu Ser Tyr Cys Cys Glu Gln Trp Lys Phe Lys  
 50 55 60

Gly Thr Pro Ser Pro Ala Leu Leu Asn Leu Gly Thr Gln Pro Lys Lys  
 65 70 75 80

55 Asp Lys Lys Leu Glu Asp Ser Ile Ala Thr Gln Leu Arg Glu Leu Pro  
 85 90 95

Glu Lys Asn Ser Asn Xaa  
 100

60

(2) INFORMATION FOR SEQ ID NO: 196:

5 (i) SEQUENCE CHARACTERISTICS:  
      (A) LENGTH: 45 amino acids  
      (B) TYPE: amino acid  
      (D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

10 Met Ala Leu Thr Phe Leu Leu Val Leu Leu Thr Leu Ala Thr Ser Ala  
1 5 10 15

His Gly Cys Thr Glu Thr Ser Asp Ala Gly Arg Ala Ser Thr Gly Gly  
20 25 30

15 Pro Gln Arg Thr Ala Arg Thr Gln Trp Leu Leu Cys Xaa  
35 40 45

20 (2) INFORMATION FOR SEQ ID NO: 197:

Gly Pro Leu Gln Gln Gly Gln His His Leu Val Glu Tyr Met Glu Arg  
20 25 30

35 Arg Leu Ala Ala Leu Glu Glu Arg Leu Ala Gln Cys Gln Asp Gln Ser  
35 40 45

Ser Arg His Ala Ala Glu Leu Arg Asp Phe Lys Asn Lys Met Leu Pro  
50 55 60

40 Leu Leu Glu Val Ala Glu Lys Glu Arg Glu Ala Leu Arg Thr Glu Ala  
65 70 75 80

Asp Thr Ile Ser Gly Arg Val Asp Arg Leu Glu Arg Glu Val Asp Tyr  
45 85 90 95

Leu Glu Thr Gln Asn Pro Ala Leu Pro Cys Val Glu Phe Asp Glu Lys  
                   100                 105                 110

50 Val Thr Gly Gly Pro Gly Thr Lys Gly Lys Gly Arg Arg Asn Glu Lys  
115 120 125

Tyr Asp Met Val Thr Asp Cys Gly Tyr Thr Ile Ser Gln Val Arg Ser  
 130 135 140

Met Lys Ile Leu Lys Arg Phe Gly Gly Pro Ala Gly Leu Trp Thr Lys  
 145 150 155 160

Asp Pro Leu Gly Gln Thr Glu Lys Ile Tyr Val Leu Asp Gly Thr Gln  
 60 165 170 175

40 (2) INFORMATION FOR SEQ ID NO: 198:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 amino acids
- (B) TYPE: amino acid

(D) TOPOLOGY: linear

Met Val Leu Pro Leu Leu Ile Phe Val	Leu Leu Pro Lys Val Val Asn		
1	5	10	15
Thr Ser Asp Pro Asp Met Arg Arg Glu	Met Glu Gln Ser Met Asn Met		
20	25	30	
Leu Asn Ser Asn His Glu Leu Pro Asp Val	Ser Glu Phe Met Thr Arg		
35	40	45	
Leu Phe Ser Ser Lys Ser Ser Gly Lys	Ser Ser Ser Gly Ser Ser Lys		
50	55	60	
Thr Gly Lys Ser Gly Ala Gly Lys Arg	Arg		

65 70

## 5 (2) INFORMATION FOR SEQ ID NO: 199:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

Met Phe Thr Met Leu Cys Ile Asn Gly Thr Thr Pro Arg Pro Leu Pro  
1 5 10 15

15 Val Pro Ser Pro Phe Gly Cys Met Ile Phe Phe Phe Lys Asn Pro  
20 25 30

20 Trp Lys Gln Arg Leu Leu Gln Gly Trp Leu Gly Ala Arg Pro Ile His  
35 40 45

Leu Leu Gly Tyr Leu Pro Leu Ser Leu Leu Trp Cys Pro Phe Pro Leu  
50 55 60

25 Pro Cys Ala Arg Cys Ser Val Val Tyr Ile Ser Ser Pro Arg His Gly  
65 70 75 80

Ala His Ala Pro Arg Asp Met Ile Leu Ser Leu Val Leu Ala His Gly  
85 90 95

30 Ala Leu Tyr Lys Glu Leu Gly Gly Arg Gly Arg Lys Trp Glu Pro Ser  
100 105 110

Xaa

35

## (2) INFORMATION FOR SEQ ID NO: 200:

40

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

Met Ala Cys Arg Cys Leu Ser Phe Leu Leu Met Gly Thr Phe Leu Ser  
1 5 10 15

50 Val Ser Gln Thr Val Leu Ala Gln Leu Asp Ala Leu Leu Val Phe Pro  
20 25 30

Gly Gln Val Ala Gln Leu Ser Cys Thr Leu Ser Pro Gln His Val Thr  
35 40 45

55 Ile Arg Asp Tyr Gly Val Ser Trp Tyr Gln Gln Arg Ala Gly Ser Ala  
50 55 60

60 Pro Arg Tyr Leu Leu Tyr Arg Ser Glu Glu Asp His His Arg Pro  
65 70 75 80

Ala Asp Ile Pro Asp Arg Phe Ser Ala Ala Lys Asp Glu Ala His Asn  
 85 90 95

5 Ala Cys Val Leu Thr Ile Ser Pro Val Gln Pro Glu Asp Asp Ala Asp  
 100 105 110

Tyr Tyr Cys Ser Val Gly Tyr Gly Phe Ser Pro  
 115 120

10

## (2) INFORMATION FOR SEQ ID NO: 201:

## 15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

20 Met Ala Gly Gly Arg Cys Gly Pro Xaa Leu Thr Ala Leu Leu Ala Ala  
 1 5 10 15

25 Trp Ile Ala Ala Val Ala Ala Thr Ala Gly Pro Glu Glu Ala Ala Leu  
 20 25 30

Pro Pro Glu Gln Ser Arg Val Gln Pro Met Thr Ala Ser Asn Trp Thr  
 35 40 45

30 Leu Val Met Glu Gly Glu Trp Met Leu Lys Phe Tyr Ala Pro Trp Cys  
 50 55 60

Pro Ser Cys Gln Gln Thr Asp Ser Glu Trp Glu Ala Phe Ala Lys Asn  
 65 70 75 80

35 Gly Glu Ile Leu Gln Ile Ser Val Gly Lys Val Asp Val Ile Gln Glu  
 85 90 95

40 Pro Gly Leu Ser Gly Arg Phe Phe Val Thr Thr Leu Pro Ala Phe Phe  
 100 105 110

His Ala Lys Asp Gly Ile Phe Arg Arg Tyr Arg Gly Pro Gly Ile Phe  
 115 120 125

45 Glu Asp Leu Gln Asn Tyr Ile Leu Glu Lys Lys Trp Gln Ser Val Glu  
 130 135 140

50 Pro Leu Thr Gly Trp Lys Ser Pro Ala Ser Leu Thr Met Ser Gly Met  
 145 150 155 160

Ala Gly Leu Phe Ser Ile Ser Gly Lys Ile Trp His Leu His Asn Tyr  
 165 170 175

55 Phe Thr Val Thr Leu Gly Ile Pro Ala Trp Cys Ser Tyr Val Phe Phe  
 180 185 190

Val Ile Ala Thr Leu Val Phe Gly Leu Phe Met Gly Leu Val Leu Val  
 195 200 205

60 Val Ile Ser Glu Cys Phe Tyr Val Pro Leu Pro Arg His Leu Ser Glu

	210	215	220	
	Arg Ser Glu Gln Asn Arg Arg Ser Glu Glu Ala His Arg Ala Glu Gln			
	225	230	235	240
5	Leu Gln Asp Ala Glu Glu Lys Asp Asp Ser Asn Glu Glu Glu Asn			
	245	250	255	
	Lys Asp Ser Leu Val Asp Asp Glu Glu Lys Glu Asp Leu Gly Asp			
10	260	265	270	
	Glu Asp Glu Ala Glu Glu Glu Asp Asn Leu Ala Ala Gly			
	275	280	285	
15	Val Asp Glu Glu Arg Ser Glu Ala Asn Asp Gln Gly Pro Pro Gly Glu			
	290	295	300	
	Asp Gly Val Thr Arg Glu Xaa Ser Arg Ala Xaa			
	305	310	315	
20				

## (2) INFORMATION FOR SEQ ID NO: 202:

25	(i) SEQUENCE CHARACTERISTICS:			
	(A) LENGTH: 236 amino acids			
	(B) TYPE: amino acid			
	(D) TOPOLOGY: linear			
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:			
30	Met Gly Thr Ala Asp Ser Asp Glu Met Ala Pro Glu Ala Pro Gln His			
	1	5	10	15
35	Thr His Ile Asp Val His Ile His Gln Glu Ser Ala Leu Ala Lys Leu			
	20	25	30	
	Leu Leu Thr Cys Cys Ser Ala Leu Arg Pro Arg Ala Thr Gln Ala Arg			
	35	40	45	
40	Gly Ser Ser Arg Leu Leu Val Ala Ser Trp Val Met Gln Ile Val Leu			
	50	55	60	
	Gly Ile Leu Ser Ala Val Leu Gly Gly Phe Phe Tyr Ile Arg Asp Tyr			
	65	70	75	80
45	Thr Leu Leu Val Thr Ser Gly Ala Ala Ile Trp Thr Gly Ala Val Ala			
	85	90	95	
50	Val Leu Ala Gly Ala Ala Phe Ile Tyr Glu Lys Arg Gly Gly Thr			
	100	105	110	
	Tyr Trp Ala Leu Leu Arg Thr Leu Leu Ala Leu Ala Phe Ser Thr			
	115	120	125	
55	Ala Ile Ala Ala Leu Lys Leu Trp Asn Glu Asp Phe Arg Tyr Gly Tyr			
	130	135	140	
	Ser Tyr Tyr Asn Ser Ala Cys Arg Ile Ser Ser Ser Asp Trp Asn			
	145	150	155	160
60				

323

Thr Pro Ala Pro Thr Gln Ser Pro Glu Glu Val Arg Arg Leu His Leu  
165 170 175

Ala Met Leu Leu Gly Val Trp Ile Leu Leu Leu Leu Ala Ser Leu Ala  
195 200 205

10 Pro Leu Trp Leu Tyr Cys Trp Arg Met Phe Pro Thr Lys Gly Lys Arg  
210 215 220

Asp Gln Lys Glu Met Leu Glu Val Ser Gly Ile Xaa  
225 230 235

15

(2) INFORMATION FOR SEQ ID NO: 203:

**20** (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

25 Met Ile His Leu Gly His Ile Leu Phe Leu Leu Leu Leu Pro Val Ala  
1 5 10 15

30 Ala Ala Gln Thr Thr Pro Gly Glu Arg Ser Ser Leu Pro Ala Phe Tyr  
30

Pro Gly Thr Ser Gly Ser Cys Ser Gly Cys Gly Ser Leu Ser Leu Pro  
35 40 45

35 Leu Leu Ala Gly Leu Val Ala Ala Asp Ala Val Ala Ser Leu Leu Ile  
50 55 60

Val Gly Ala Val Phe Leu Cys Ala Arg Pro Arg Arg Ser Pro Ala Gln  
65 70 75 80

40 Glu Asp Gly Lys Val Tyr Ile Asn Met Pro Gly Arg Gly  
85 90

45

(2) INFORMATION FOR SEQ ID NO: 204:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 204:

55 Met Trp Ser Ala Gly Arg Gly Gly Ala Ala Trp Pro Val Leu Leu Gly  
1 5 10 15

Leu Leu Leu Ala Leu Leu Val Pro Gly Gly Gly Ala Ala Lys Thr Gly  
20 25 30

60 Ala Asp Ser

35

## 5 (2) INFORMATION FOR SEQ ID NO: 205:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

Asp Cys Xaa His Val Ser Val Leu Gln Ser Thr Ile Ser Pro Leu Leu  
1 5 10 15

15 Pro Leu Pro Leu Leu Leu Pro His Gly Asn Cys Glu Glu Ala Pro Trp  
20 25 30

Gln Ala Ala Val Ile Gly Gly Asp Arg Ile  
20 35 40

## 25 (2) INFORMATION FOR SEQ ID NO: 206:

## 25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

Met Arg Asp Cys Leu Ser Leu Lys Pro Arg Pro Leu Phe Pro Thr Gln  
1 5 10 15

35 Phe Phe Phe Ile Leu Leu Ile Phe Ile Ala Glu Val Ala Ala Ala  
20 25 30

Val Val Ala Leu Val Tyr Thr Thr Met Val Arg His Trp Asp Gly Gly  
35 40 45

40 Arg Glu Glu Asp Trp Ala Lys Pro Trp Glu Trp Ala Val Ala Cys Glu  
50 55 60

45 Trp Pro Pro Ser Val Pro Ala Pro Lys His Trp Pro Ala Ser Pro Arg  
65 70 75 80

Leu Ser Thr Ser Xaa  
85

50

## (2) INFORMATION FOR SEQ ID NO: 207:

## 55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

60 Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Met

325

1	5	10	15
Gln Phe Leu Cys His Glu Phe Leu Arg Xaa Asn Pro Arg Val Thr Arg			
20                            25                            30			
5	Leu Leu Ser Glu Met Arg Ile His Leu Leu Pro Ser Met Asn Pro Asp		
35                            40                            45			
10	Gly Tyr Glu Ile Ala Tyr His Arg Gly Ser Glu Leu Val Gly Trp Ala		
50                            55                            60			
Glu Gly Arg Trp Asn Asn Gln Ser Ile Asp Leu Asn His Asn Phe Ala			
65                            70                            75                            80			
15	Xaa Leu Asn Thr Pro Leu Trp Glu Ala Gln Asp Asp Gly Lys Val Pro		
85                            90                            95			
His Ile Val Pro Asn His His Leu Pro Leu Pro Thr Tyr Tyr Thr Leu			
100                          105                          110			
20	Pro Asn Ala Thr Val Ala Pro Glu Thr Arg Ala Val Ile Lys Trp Met		
115                          120                          125			
25	Lys Arg Ile Pro Phe Val Leu Ser Ala Asn Leu His Gly Gly Glu Leu		
130                          135                          140			
Val Val Ser Tyr Pro Phe Asp Met Thr Arg Thr Pro Trp Ala Ala Arg			
145                          150                          155                          160			
30	Glu Leu Thr Pro Thr Pro Asp Asp Ala Val Phe Arg Trp Leu Ser Thr		
165                          170                          175			
Val Tyr Ala Gly Ser Asn Leu Ala Met Gln Asp Thr Ser Arg Arg Pro			
180                          185                          190			
35	Cys His Ser Gln Asp Phe Ser Val His Gly Asn Ile Ile Asn Gly Ala		
195                          200                          205			
40			

## (2) INFORMATION FOR SEQ ID NO: 208:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

Met	Glu	Ile	Ser	Cys	Leu	Leu	Leu	Ile	Gln	Asp	Ser	Asp	Glu	Met
1	5				10				15					

55 Glu Asp Gly Pro Gly Val Gln Asp  
 20

## 60 (2) INFORMATION FOR SEQ ID NO: 209:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 483 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

Met Ala Thr Gly Gly Gly Ile Arg Ala Met Thr Ser Leu Tyr Gly Gln			
1	5	10	15

Leu Ala Gly Leu Lys Glu Leu Gly Leu Leu Asp Cys Xaa Ser Tyr Ile			
20	25	30	

Thr Gly Ala Ser Gly Ser Thr Trp Ala Leu Ala Asn Leu Tyr Lys Asp			
35	40	45	

Pro Glu Trp Ser Gln Lys Asp Leu Ala Gly Pro Thr Glu Leu Leu Lys			
50	55	60	

Thr Gln Val Thr Lys Asn Lys Leu Gly Val Leu Ala Pro Ser Gln Leu			
65	70	75	80

Gln Arg Tyr Arg Gln Glu Leu Ala Glu Arg Ala Arg Leu Gly Tyr Pro			
85	90	95	

25 Ser Cys Phe Thr Asn Leu Trp Ala Leu Ile Asn Glu Ala Leu Leu His			
100	105	110	

Asp Glu Pro His Asp His Lys Leu Ser Asp Gln Arg Glu Ala Leu Ser			
115	120	125	

His Gly Gln Asn Pro Leu Pro Ile Tyr Cys Ala Leu Asn Thr Lys Gly			
130	135	140	

35 Gln Ser Leu Thr Thr Phe Glu Phe Gly Glu Trp Cys Glu Phe Ser Pro			
145	150	155	160

Tyr Glu Val Gly Phe Pro Lys Tyr Gly Ala Phe Ile Pro Ser Glu Leu			
165	170	175	

40 Phe Gly Ser Glu Phe Phe Met Gly Gln Leu Met Lys Arg Leu Pro Glu			
180	185	190	

Ser Arg Ile Cys Phe Leu Glu Gly Ile Trp Ser Asn Leu Tyr Ala Ala			
45 195	200	205	

Asn Leu Gln Asp Ser Leu Tyr Trp Ala Ser Glu Pro Ser Gln Phe Trp			
210	215	220	

50 Asp Arg Trp Val Arg Asn Gln Ala Asn Leu Asp Lys Glu Gln Val Pro			
225	230	235	240

Leu Leu Lys Ile Glu Glu Pro Pro Ser Thr Ala Gly Arg Ile Ala Glu			
245	250	255	

55 Phe Phe Thr Asp Leu Leu Thr Trp Arg Pro Leu Ala Gln Ala Thr His			
260	265	270	

Asn Phe Leu Arg Gly Leu His Phe His Lys Asp Tyr Phe Gln His Pro			
60 275	280	285	

His Phe Ser Thr Trp Lys Ala Thr Thr Leu Asp Gly Leu Pro Asn Gln  
290 295 300

5 Leu Thr Pro Ser Glu Pro His Leu Cys Leu Leu Asp Val Gly Tyr Leu  
305 310 315 320

Ile Asn Thr Ser Cys Leu Pro Leu Leu Gln Pro Thr Arg Asp Val Asp  
325 330 335

10 Leu Ile Leu Ser Leu Asp Tyr Asn Leu His Gly Ala Phe Gln Gln Leu  
340 345 350

Gln Leu Leu Gly Arg Phe Cys Gln Glu Gln Gly Ile Pro Phe Pro Pro  
15 355 360 365

Ile Ser Pro Ser Pro Glu Glu Gln Leu Gln Pro Arg Glu Cys His Thr  
370 375 380

20 Phe Ser Asp Pro Thr Cys Pro Gly Ala Pro Ala Val Leu His Phe Pro  
385 390 395 400

Leu Val Ser Asp Ser Phe Arg Glu Tyr Ser Ala Pro Gly Val Arg Arg  
405 410 415

25 Thr Pro Glu Glu Ala Ala Ala Gly Glu Val Asn Leu Ser Ser Ser Asp  
420 425 430

Ser Pro Tyr His Tyr Thr Lys Val Thr Tyr Ser Gln Glu Asp Val Asp  
30 435 440 445

Lys Leu Leu His Leu Thr His Tyr Asn Val Cys Asn Asn Gln Glu Gln  
450 455 460

35 Leu Leu Glu Ala Leu Arg Gln Ala Val Gln Arg Arg Arg Gln Arg Arg  
465 470 475 480

Pro His Xaa

40

## (2) INFORMATION FOR SEQ ID NO: 210:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 13 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

50 Leu Glu Val Gly Cys Ile Gln Val Ala Pro Asp Thr Phe  
1 5 10

55 (2) INFORMATION FOR SEQ ID NO: 211:

- 60 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

Met Ser Leu Phe Phe Leu Leu Thr Leu Ile Ser Lys Leu His Gly Asp  
5 1 5 10 15  
Ala Glu Val Cys  
20

10

(2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 55 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:  
  
20 Met Pro His Pro Pro Leu Pro Glu Thr Ser Leu Glu Ala Gln Leu Pro  
1 5 10 15  
Met Gly Leu Leu Gln Leu Leu Arg Cys Ser Val Gln Ala Trp Ser Pro  
20 25 30  
25 Pro Pro Ser Ser Phe Cys Pro Gly Ser Glu Pro Arg Ser Ala Ser Ala  
35 40 45  
His Trp Gly Tyr Trp Trp Pro  
30 50 55

(2) INFORMATION FOR SEQ ID NO: 213:

35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:  
  
Asp Pro Glu Thr Arg Trp His His Gly Gly Ser Ala Gln Asn Gly Leu  
1 5 10 15  
45 Leu Met Leu Ile Ser Val Leu Gln Gln Pro Val Ile Gly Thr Gly Ser  
20 25 30  
Tyr Leu Cys  
35  
50

(2) INFORMATION FOR SEQ ID NO: 214:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 230 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:  
60

Met Glu Pro Leu Arg Leu Leu Ile Leu Phe Val Thr Glu Leu Ser  
 1 5 10 15

Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu  
 5 20 25 30

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys  
 35 40 45

10 Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val  
 50 55 60

Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly  
 65 70 75 80

15 Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr  
 85 90 95

20 Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser  
 100 105 110

Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val  
 115 120 125

25 Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro  
 130 135 140

Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser  
 145 150 155 160

30 Arg Ser Leu Leu Glu Gly Glu Ile Pro Phe Pro Pro Thr Ser Ile Leu  
 165 170 175

Leu Leu Leu Ala Cys Ile Phe Leu Ile Lys Ile Leu Ala Ala Ser Xaa  
 35 180 185 190

Leu Trp Ala Ala Ala Trp His Gly Gln Lys Pro Gly Thr His Pro Pro  
 195 200 205

40 Ser Glu Leu Asp Cys Gly His Asp Pro Gly Tyr Gln Leu Gln Thr Leu  
 210 215 220

Pro Gly Leu Arg Asp Thr  
 225 230

45

## (2) INFORMATION FOR SEQ ID NO: 215:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 231 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

55 Met Glu Pro Leu Arg Leu Leu Ile Leu Phe Val Thr Glu Leu Ser  
 1 5 10 15

Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu  
 60 20 25 30

330

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys  
 35 40 45  
 5 Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val  
 50 55 60  
 Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly  
 65 70 75 80  
 10 Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr  
 85 90 95  
 Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser  
 15 100 105 110  
 Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val  
 115 120 125  
 20 Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro  
 130 135 140  
 Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser  
 145 150 155 160  
 25 Arg Ser Leu Leu Glu Gly Glu Ile Pro Phe Pro Pro Thr Ser Ile Leu  
 165 170 175  
 Leu Leu Leu Ala Cys Ile Phe Leu Ile Lys Ile Leu Ala Ala Ser Ala  
 30 180 185 190  
 Leu Trp Ala Ala Ala Trp His Gly Gln Lys Pro Gly Thr His Pro Pro  
 195 200 205  
 35 Ser Glu Leu Asp Cys Gly His Asp Pro Gly Tyr Gln Leu Gln Thr Leu  
 210 215 220  
 Pro Gly Leu Arg Asp Thr Xaa  
 225 230  
 40

## (2) INFORMATION FOR SEQ ID NO: 216:

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 127 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:  
 50 Met Gly Leu Thr Gly Phe Gly Val Phe Phe Phe Gly Met Ile  
 1 5 10 15  
 Leu Phe Phe Asp Lys Ala Leu Ala Ile Gly Asn Val Leu Phe Val  
 55 20 25 30  
 Ala Gly Leu Ala Phe Val Ile Gly Leu Glu Arg Thr Phe Arg Phe Phe  
 35 40 45  
 60 Phe Gln Lys His Lys Met Lys Ala Thr Gly Phe Phe Leu Gly Val

50	55	60	
Phe Val Val Leu Ile Gly Trp Pro Leu Ile Gly Met Ile Phe Glu Ile			
65	70	75	80
Tyr Gly Phe Phe Leu Leu Phe Arg Gly Phe Phe Pro Val Val Gly			
5	85	90	95
Phe Ile Arg Arg Val Pro Val Leu Gly Ser Leu Leu Asn Leu Pro Gly			
10	100	105	110
Ile Arg Ser Phe Val Asp Lys Val Gly Glu Ser Asn Asn Met Val			
	115	120	125

15

(2) INFORMATION FOR SEQ ID NO: 217:

(i) SEQUENCE CHARACTERISTICS:

20	(A) LENGTH: 47 amino acids			
	(B) TYPE: amino acid			
	(D) TOPOLOGY: linear			
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:				
25	Met Ile Arg Lys Leu His Lys Ile Ile Val Phe Ser Pro Arg Val Ile			
	1	5	10	15
Val Leu Leu Asn Cys Phe Phe Ile Lys Ala Lys Phe Val Leu Tyr				
	20	25	30	
30	Ile Phe Val Phe His Val Leu Asp Gly Ser Ile Ser Tyr Pro Val			
	35	40	45	

35

(2) INFORMATION FOR SEQ ID NO: 218:

(i) SEQUENCE CHARACTERISTICS:

40	(A) LENGTH: 41 amino acids			
	(B) TYPE: amino acid			
	(D) TOPOLOGY: linear			
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:				
45	Met Leu Leu Asn Gln His Phe Lys Ile Phe Gly Ser Leu Ile His Met			
	1	5	10	15
Asn Leu Leu Phe Ala Leu Ile Ser Leu Gly Ser Ser Asn Leu Ser Gly				
	20	25	30	
50	Val Gln Phe Cys Cys Glu Thr Val Gln			
	35	40		

55

(2) INFORMATION FOR SEQ ID NO: 219:

(i) SEQUENCE CHARACTERISTICS:

60	(A) LENGTH: 105 amino acids	
	(B) TYPE: amino acid	
	(D) TOPOLOGY: linear	

332

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

Met Gln Pro Leu Asn Phe Ser Ser Thr Xaa Cys Ser Ser Phe Ser Pro  
1 5 10 15

5 Pro Thr Thr Val Ile Leu Leu Ile Leu Leu Cys Phe Glu Gly Leu Leu  
20 25 30

Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile  
10 35 40 45

Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg  
50 55 60

15 Trp Ala Lys Lys Thr Lys Trp Met Asn Met Lys Ala Val Phe Gly His  
65 70 75 80

Pro Phe Ser Leu Gly Trp Ala Ser Pro Phe Ala Thr Pro Asp Gln Gly  
85 90 95

20 Lys Ala Asp Pro Tyr Gln Tyr Val Val  
100 105

25

## (2) INFORMATION FOR SEQ ID NO: 220:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 amino acids  
30 (B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

Met Tyr Thr Asn His Phe Asn Leu Tyr Leu Lys Tyr Ile Leu Leu Ile  
35 1 5 10 15

Ile Leu Ile Leu Asn Met Thr Asn Ser Ser Ser Arg Tyr  
20 25

40

## (2) INFORMATION FOR SEQ ID NO: 221:

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

50 Met Asn Glu Leu Leu Phe Phe Phe Phe Phe Phe Leu His Phe  
1 5 10 15

Val

55

## (2) INFORMATION FOR SEQ ID NO: 222:

60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

5

Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu  
1 5 10 15

10

Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala  
20 25 30

Gln Glu Ala Gly Thr Ser Lys Pro Asn Glu Glu Ile Ser Gly Pro Ala  
35 40 45

15

Glu Pro Ala Ser Pro Pro Glu Thr Thr Thr Ala Gln Glu Xaa Ser  
50 55 60

Ala Ala Ala Val Gln Gly Thr Ala Lys Val Thr Ser Ser Arg Gln Glu  
65 70 75 80

20

Leu Asn Pro Leu Lys Ser Ile Val Glu Lys Ser Ile Leu Leu Thr Glu  
85 90 95

25

Gln Ala Leu Ala Lys Ala Gly Lys Gly Met His Gly Gly Val Pro Gly  
100 105 110

Gly Lys Gln Phe Ile Glu Asn Gly Ser Glu Phe Ala Gln Lys Leu Leu  
115 120 125

30

Lys Lys Phe Ser Leu Leu Lys Pro Trp Ala  
130 135

35

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

40

Met Leu Gly Cys Gly Ile Pro Ala Leu Gly Leu Leu Leu Leu Gln  
1 5 10 15

45

Xaa Ser Ala Asp Gly Asn Gly Ile Gln Gly Phe Phe Tyr Pro Trp Ser  
20 25 30

50

Cys Glu Gly Asp Ile Trp Asp Arg Glu Ser Cys Gly Gly Gln Ala Ala  
35 40 45

Ile Arg  
50

55

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

5 Met Glu Ala Val Phe Thr Val Phe Phe Phe Leu Leu Phe Cys Phe  
1 5 10 15

10 (2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 155 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

	Met	Gly	Phe	Gly	Ala	Thr	Leu	Ala	Val	Gly	Leu	Thr	Ile	Phe	Val	Leu
20	1				5					10					15	
	Ser	Val	Val	Thr	Ile	Ile	Ile	Cys	Phe	Thr	Cys	Ser	Cys	Cys	Cys	Leu
					20					25					30	
25	Tyr	Lys	Thr	Cys	Arg	Arg	Pro	Arg	Pro	Val	Val	Thr	Thr	Thr	Thr	Ser
					35					40					45	
	Thr	Thr	Val	Val	His	Ala	Pro	Tyr	Pro	Gln	Pro	Pro	Ser	Val	Pro	Pro
					50					55					60	
30	Ser	Tyr	Pro	Gly	Pro	Ser	Tyr	Gln	Gly	Tyr	His	Thr	Met	Pro	Pro	Gln
					65					70			75			80
	Pro	Gly	Met	Pro	Ala	Ala	Pro	Tyr	Pro	Met	Gln	Tyr	Pro	Pro	Pro	Tyr
							85					90				95
35	Pro	Ala	Gln	Pro	Met	Gly	Pro	Pro	Ala	Tyr	His	Glu	Thr	Leu	Ala	Gly
							100					105				110
	Gly	Ala	Ala	Ala	Pro	Tyr	Pro	Ala	Ser	Gln	Pro	Pro	Tyr	Asn	Pro	Xaa
40																
					115					120					125	
	Tyr	Met	Asp	Ala	Pro	Lys	Xaa	Xaa	Ser	Glu	His	Ser	Leu	Ala	Ser	Leu
							130						140			
45	Ala	Ala	Thr	Trp	Leu	Cys	Cys	Val	Cys	Ala	Xaa					
								145				155				

50 (2) INFORMATION FOR SEQ ID NO: 226:

Met Gly Phe Gly Ala Thr Leu Ala Val Gly  
1 5 10

## (2) INFORMATION FOR SEQ ID NO: 227:

## 5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

10 Met Ser Ile Phe Leu Val Met Ser Ile Ser Cys Ser Ser Thr Ser His  
1 5 10 15

Cys Tyr Ser Phe  
15 20

## (2) INFORMATION FOR SEQ ID NO: 228:

## 20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile  
1 5 10 15

Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys  
20 25 30

Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val  
35 40 45

Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu  
50 55 60

Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser  
65 70 75 80

Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr Xaa  
85 90

45

## (2) INFORMATION FOR SEQ ID NO: 229:

## 50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

55 Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile  
1 5 10 15

Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys  
20 25 30

60

Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val  
35 40 45

Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu  
5 50 55 60

Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser  
65 70 75 80

10 Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr Xaa  
85 90

## 15 (2) INFORMATION FOR SEQ ID NO: 230:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37 amino acids  
(B) TYPE: amino acid  
20 (D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

Met Gly Trp Ser Ala Gly Leu Leu Phe Leu Leu Ile Leu Tyr Leu Pro  
1 5 10 15

25 Val Pro Gly Trp Met Glu Arg Glu Asp Gly Gly Asp Gly Thr Ser Phe  
20 25 30

30 Thr Ser Gly Ser Trp  
35

35 (2) INFORMATION FOR SEQ ID NO: 231:  
35

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 81 amino acids  
(B) TYPE: amino acid  
40 (D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser  
1 5 10 15

45 Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Leu Ala His Val Gln Thr  
20 25 30

Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro  
35 40 45

50 Ile Lys Lys Ile Leu Gly Ile Phe Ile Ile Arg Thr Tyr Leu Arg Lys  
50 55 60

55 Ile Val Ile Ala Phe Met Leu Trp Ser Pro Cys Leu Cys Gly Gly Leu  
65 70 75 80

Met

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 301 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

10 Met Asp Ala Arg Trp Trp Ala Val Val Val Leu Ala Ala Phe Pro Ser  
1 5 10 15

Leu Gly Ala Gly Gly Glu Thr Pro Glu Ala Pro Pro Pro Glu Ser Trp Thr  
                   20                         25                         30

15 Gln Leu Trp Phe Phe Arg Phe Val Val Asn Ala Ala Gly Tyr Ala Xaa  
           35                  40                  45

Phe Met Val Pro Gly Tyr Leu Leu Val Gln Tyr Phe Arg Arg Lys Asn  
50 55 60

Tyr Leu Glu Thr Gly Arg Gly Leu Cys Phe Pro Leu Val Lys Ala Cys  
50 55 60 65 70 75 80

Tyr Leu Glu Thr Gly Arg Gly Leu Cys Phe Pro Leu Val

Arg Thr Glu Ala Ala Glu Thr Thr Pro Met Trp Gln Ala Leu Lys Leu  
100 105 110

30 Leu Phe Cys Ala Thr Gly Leu Gln Val Ser Tyr Leu Thr Trp Gly Val  
           115               120               125

Leu Gln Glu Arg Val Met Thr Arg Ser Tyr Gly Ala Thr Ala Thr Ser  
130 135 140

Pro Gly Glu Arg Phe Thr Asp Ser Gln Phe Leu Val Leu Met Asn Arg

Pro Gly Glu Arg Phe Thr Asp Ser Gln Phe Leu Val Leu

40 Val Leu Ala Leu Ile Val Ala Gly Leu Ser Cys Val Leu Cys Lys Gln

165 170 175

Pro Arg His Gly Ala Pro Met Tyr Arg Tyr Ser Phe Ala Ser Leu Ser  
180 185 190

Arg Val Ileu Ser Ser Thr Cys Glu Tyr Glu Ala Leu Lys Phe

50 Phe Pro Thr Gln Val Leu Ala Lys Ala Ser Lys Val Ile Pro Val Met  
210 215 220

Leu Met Gly Lys Leu Val Ser Arg Arg Xaa Asn Glu His Trp Glu Tyr  
225 230 235 240

55 Leu Thr Ala Thr Leu Ile Ser Ile Gly Val Ser Met Phe Leu Leu Ser  
245 250 255

Ser Gly Pro Glu Pro Arg Ser Ser Pro Ala Thr Thr Leu Ser Gly Leu  
260 265 270

Ile Leu Leu Ala Gly Tyr Ile Ala Phe Asp Ser Phe Thr Ser Asn Trp  
275 280 285

5 Gln Asp Ala Cys Leu Pro Ile Arg Cys His Arg Cys Arg  
 290 295 300

(2) INFORMATION FOR SEQ ID NO: 233:

10

#### **INFLUENCE CHARACTERISTICS:**

(A) LENGTH: 313 amino

(B) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

Met Ser Asp Leu Leu Leu Leu Gly Leu Ile Gly Gly Leu Thr Leu Leu  
1 5 10 15

20 Leu Leu Leu Thr Leu Leu Ala Phe Ala Gly Tyr Ser Gly Leu Leu Ala  
20 25 30

Gly Val Glu Val Ser Ala Gly Ser Pro Pro Ile Arg Asn Val Thr Val  
35 40 45

Thr Glu Ser Cys Ser Ile Ser Pro Lys Leu Arg Ser Ile Ala Val Tyr  
 65                    70                    75                    80

Tyr Asp Asn Pro His Met Val Pro Pro Asp Lys Cys Arg Cys Ala Val  
85 90 95

Gly Ser Ile Leu Ser Glu Gly Glu Ser Pro Ser Pro Glu Leu Ile  
100 105 110

35 Gly Ser Ile Leu Ser Glu Gly Glu Glu Ser Pro Ser Pro Glu Leu Ile  
100 105 110

Asp Leu Tyr Gln Lys Phe Gly Phe Lys Val Phe Ser Phe Pro Ala Pro  
115 120 125

Trp Leu Ala Thr Arg Arg Val His Pro Ala Leu Asp Thr Tyr Ile Lys  
 145                    150                    155                    160

Glu Arg Lys Leu Cys Ala Tyr Pro Arg Leu Glu Ile Tyr Gln Glu Asp

Gln Ile His Phe Met Cys Pro Leu Ala Xaa Gln Gly Asp Phe Tyr Val

50 Gln Ile His Phe Met Cys Pro Leu Ala Xaa Gln Gly Asp Phe Tyr Val  
180 185 190

Pro Glu Met Lys Glu Thr Glu Trp Lys Trp Arg Gly Leu Val Glu Ala  
195 200 205

55 Ile Asp Thr Gln Val Asp Gly Thr Gly Ala Asp Thr Met Ser Asp Thr  
           210               215               220

Ser Ser Val Ser Leu Glu Val Ser Pro Gly Ser Arg Glu Thr Ser Ala  
225 230 235 240

Ala Thr Leu Ser Pro Gly Ala Ser Ser Arg Gly Trp Asp Asp Gly Asp  
245 250 255

5 Thr Arg Ser Glu His Ser Tyr Ser Glu Ser Gly Ala Ser Gly Ser Ser  
260 265 270

Phe Glu Glu Leu Asp Leu Glu Gly Glu Gly Pro Leu Gly Glu Ser Arg  
275 280 285

10 Leu Asp Pro Gly Thr Xaa Pro Leu Gly Thr Thr Lys Trp Leu Trp Glu  
290 295 300

Pro Thr Ala Pro Glu Lys Gly Lys Glu  
15 305 310

(2) INFORMATION FOR SEQ ID NO: 234:

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 48 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

Pro Gln Ser Leu Ile Leu His Leu Leu Leu Phe Phe Phe Leu Leu Phe  
1 5 10 15

30 Leu Phe Phe Ile Phe Ile Phe Leu Phe Phe Leu Gln Cys Leu Thr Phe  
20 25 30

Leu Phe Xaa Lys Pro Arg Gly Arg Tyr His Gly Leu Cys Phe Lys Phe  
35 40 45

35

40 (2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp Leu  
50 1 5 10 15

Cys Cys Ala Thr Pro Arg Met His Cys Ser Val Glu Met Ala Met Asn  
20 25 30

55 Pro Val

60 (2) INFORMATION FOR SEQ ID NO: 236:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 313 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

	Met Thr Arg Gly Gly Pro Gly Gly Arg Pro Gly Leu Pro Gln Pro Pro	
	1 5 10 15	
10	Pro Leu Leu Leu Leu Leu Xaa Leu Leu Leu Val Thr Ala Glu	
	20 25 30	
15	Pro Pro Lys Pro Ala Gly Val Tyr Tyr Ala Thr Ala Tyr Trp Met Pro	
	35 40 45	
	Ala Glu Lys Thr Val Gln Val Lys Asn Val Met Asp Lys Asn Gly Asp	
	50 55 60	
20	Ala Tyr Gly Phe Tyr Asn Asn Ser Val Lys Thr Thr Gly Trp Gly Ile	
	65 70 75 80	
	Leu Glu Ile Arg Ala Gly Tyr Gly Ser Gln Thr Leu Ser Asn Glu Ile	
	85 90 95	
25	Ile Met Phe Val Ala Gly Phe Leu Glu Gly Tyr Leu Thr Ala Pro His	
	100 105 110	
30	Met Asn Asp His Tyr Thr Asn Leu Tyr Pro Gln Leu Ile Thr Lys Pro	
	115 120 125	
	Ser Ile Met Asp Lys Val Gln Asp Phe Met Glu Lys Gln Asp Lys Trp	
	130 135 140	
35	Thr Arg Lys Asn Ile Lys Glu Tyr Lys Thr Asp Ser Phe Trp Arg His	
	145 150 155 160	
	Thr Gly Tyr Val Met Ala Gln Ile Asp Gly Leu Tyr Val Gly Ala Lys	
	165 170 175	
40	Lys Arg Ala Ile Leu Glu Gly Thr Lys Pro Met Thr Leu Phe Gln Ile	
	180 185 190	
45	Gln Phe Leu Asn Ser Val Gly Asp Leu Leu Asp Leu Ile Pro Ser Leu	
	195 200 205	
	Ser Pro Thr Lys Asn Gly Ser Leu Lys Val Phe Lys Arg Trp Asp Met	
	210 215 220	
50	Gly His Cys Ser Ala Leu Ile Lys Val Leu Pro Gly Phe Glu Asn Ile	
	225 230 235 240	
	Leu Phe Ala His Ser Ser Trp Tyr Thr Tyr Ala Ala Met Leu Arg Ile	
	245 250 255	
55	Tyr Lys His Trp Asp Phe Asn Xaa Ile Asp Lys Asp Thr Ser Ser Ser	
	260 265 270	
60	Arg Leu Ser Phe Ser Ser Tyr Pro Gly Phe Leu Glu Ser Leu Asp Asp	
	275 280 285	

Phe Tyr Ile Leu Ser Ser Gly Leu Ile Leu Leu Gln Thr Thr Asn Ser  
 290 295 300

5 Val Phe Asn Lys Thr Leu Leu Lys Gln  
 305 310

10 (2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 296 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His  
 1 5 10 15

20 Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp  
 20 25 30

25 Phe Ser Tyr Lys Arg Xaa Asn Cys Lys Pro Ile Pro Val Asn Leu Gln  
 35 40 45

Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu  
 50 55 60

30 Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile  
 65 70 75 80

Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys  
 85 90 95

35 Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln  
 100 105 110

40 Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val  
 115 120 125

Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg  
 130 135 140

45 Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His  
 145 150 155 160

Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys  
 165 170 175

50 Asn Lys Asn Asp Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn  
 180 185 190

55 Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg  
 195 200 205

Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu  
 210 215 220

60 Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys

225	230	235	240
Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro			
	245	250	255
5	Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser		
	260	265	270
10	Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg		
	275	280	285
Ser Ile Arg Lys Leu Gln Cys Xaa			
	290	295	

15

(2) INFORMATION FOR SEQ ID NO: 238:

20	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 92 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:			
25	Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr		
	1	5	10
	15		
Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp			
	20	25	30
30	Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe		
	35	40	45
Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe			
35	50	55	60
Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Gly Thr Phe			
	65	70	75
	80		
40	Arg Ser Ser Ile Arg Arg Leu Ser Xaa Arg Xaa Arg		
	85	90	

45 (2) INFORMATION FOR SEQ ID NO: 239:

50	(i) SEQUENCE CHARACTERISTICS:		
	(A) LENGTH: 71 amino acids		
	(B) TYPE: amino acid		
	(D) TOPOLOGY: linear		
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:			
55	Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Phe Leu Phe Ile		
	1	5	10
	15		
Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro			
	20	25	30
Phe Pro Pro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Ser			
60	35	40	45

Pro Pro Leu Pro Pro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro  
50 55 60

5 Ser Arg Gly Cys Val Leu Leu  
65 70

10 (2) INFORMATION FOR SEQ ID NO: 240:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 71 amino acids  
(B) TYPE: amino acid  
15 (D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Phe Ile  
1 5 10 15

20 Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro  
20 25 30

25 Phe Pro Pro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Ser  
35 40 45

Pro Pro Leu Pro Pro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro  
50 55 60

30 Ser Arg Gly Cys Val Leu Leu  
65 70

35 (2) INFORMATION FOR SEQ ID NO: 241:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 amino acids  
(B) TYPE: amino acid  
40 (D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

Met Phe Tyr Val Leu Ser Val Xaa Leu Xaa Leu Phe Leu Ala Cys  
1 5 10 15

45 Gly Leu Cys Leu Xaa Leu Leu Thr Gly Lys Leu Leu  
20 25

50 (2) INFORMATION FOR SEQ ID NO: 242:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 58 amino acids  
55 (B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

60 Met Lys Leu Phe Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly  
1 5 10 15

His Ile Leu Ala Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu  
20 25 30

5 Gly Pro Gly Trp Val Pro Ser Ala Leu Xaa Arg Leu His Pro Gly His  
35 40 45

Leu Ser Gly Ser Val Leu Val Ser Ala Ala  
50 55

10

## (2) INFORMATION FOR SEQ ID NO: 243:

## 15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

20 Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe Val  
1 5 10 15

25 Trp Ala Ala His Asn Lys Asp Val Leu Arg Arg Met Lys Lys Arg Tyr  
20 25 30

Pro Thr Thr Phe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu Ile  
35 40 45

30 Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe Pro  
50 55 60

Leu Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu Lys  
65 70 75 80

35 Asn Lys Leu Glu Asn Lys Met Glu Gly Ile Gly Leu Lys Arg Thr Pro  
85 90 95

40 Met Gly Ile Val Leu Asp Ala Leu Glu Gln Gln Glu Gly Ile Asn  
100 105 110

Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu  
115 120

45

## (2) INFORMATION FOR SEQ ID NO: 244:

## 50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

55 Ala Leu Val Ser Gly Gln Leu Cys Met Glu Ile Ala Arg Gly Asn Ile  
1 5 10 15

Phe Phe Leu Asn Xaa Leu Val Thr Thr Phe Cys Cys Ser Cys Leu Leu  
20 25 30

60



	100	105	110
	Arg His Ser Ser Met Trp Ser Leu Ser Arg Gln Thr Ser Pro Ala Ser		
	115	120	125
5	Ser Leu Thr Gly Ala Thr Phe Arg Lys Leu Asp Glu Lys Gly Ser Leu		
	130	135	140
	Gln Trp Asp Arg Ile Thr Arg Leu Glu Lys Gly Lys Ile Tyr Arg Gln		
10	145	150	155
	Gly Asn Leu Phe Asp Phe Leu Arg Leu Thr Glu Trp Arg Gly Pro Arg		
	165	170	175
15	Val Leu Tyr Phe Gly Asp His Leu Tyr Ser Asp Leu Ala Asp Leu Met		
	180	185	190
	Leu Arg His Gly Trp Arg Thr Gly Ala Ile Ile Pro Glu Leu Glu Arg		
	195	200	205
20	Glu Ile Arg Ile Ile Asn Thr Glu Gln Tyr Met His Ser Leu Thr Trp		
	210	215	220
	Gln Gln Ala Leu Thr Gly Leu Leu Glu Arg Met Gln Thr Tyr Gln Asp		
25	225	230	235
	Ala Glu Ser Arg Gln Val Leu Ala Ala Trp Met Lys Glu Arg Gln Glu		
	245	250	255
30	Leu Arg Cys Ile Thr Lys Ala Leu Phe Asn Ala Gln Phe Gly Ser Ile		
	260	265	270
	Phe Arg Thr Phe His Asn Pro Thr Tyr Phe Ser Arg Arg Leu Val Arg		
	275	280	285
35	Phe Ser Asp Leu Tyr Met Ala Ser Leu Ser Cys Leu Leu Asn Tyr Arg		
	290	295	300
	Val Asp Phe Thr Phe Tyr Pro Arg Arg Thr Pro Leu Gln His Glu Ala		
40	305	310	315
	Pro Leu Trp Met Asp Gln Leu Leu His Arg Leu His Glu Asp Pro Leu		
	325	330	335
45	Pro Trp Xaa		

50 (2) INFORMATION FOR SEQ ID NO: 247:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

Met Ala Leu Leu Ser Cys Val Val Asp Tyr Phe Leu Gly His Ser Leu			
1	5	10	15

Xaa Val

5

## (2) INFORMATION FOR SEQ ID NO: 248:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

15	Met Asn Trp Glu Leu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu	1	5	10	15
	Leu Leu Leu Val Gln Leu Leu Arg Phe Leu Arg Ala Asp Gly Asp Leu				
	20	25		30	
20	Thr Leu Leu Trp Ala Glu Trp Gln Gly Arg Arg Pro Glu Trp Glu Leu	35	40	45	
	Thr Asp Met Val Val Trp Val Thr Gly Ala Ser Ser Gly Ile Gly Glu	50	55	60	
25	Glu Leu Ala Tyr Gln Leu Ser Lys Leu Gly Val Ser Leu Val Leu Ser	65	70	75	80
	Ala Arg Arg Val His Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Glu	85	90	95	
30	Asn Gly Asn Leu Lys Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu	100	105	110	
35	Thr Asp Thr Gly Ser His Glu Ala Ala Thr Lys Ala Val Leu Gln Glu	115	120	125	
	Phe Gly Arg Ile Asp Ile Leu Val Asn Asn Gly Gly Met Ser Gln Arg	130	135	140	
40	Ser Leu Cys Met Asp Thr Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu	145	150	155	160
	Leu Asn Tyr Leu Gly Thr Val Ser Leu Thr Lys Cys Val Leu Pro His	165	170	175	
45	Met Ile Glu Arg Lys Gln Gly Lys Ile Val Thr Val Asn Ser Ile Leu	180	185	190	
50	Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala Ser Lys His	195	200	205	
	Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr	210	215	220	
55	Pro Gly Ile Ile Val Ser Asn Ile Cys Pro Gly Pro Val Gln Ser Asn	225	230	235	240
	Ile Val Glu Asn Ser Leu Ala Gly Glu Val Thr Lys Thr Ile Gly Asn	245	250	255	

Asn Gly Asp Gln Ser His Lys Met Thr Thr Ser Arg Cys Val Arg Leu  
260 265 270

5 Met Leu Ile Ser Met Ala Asn Asp Leu Lys Glu Val Trp Ile Ser Glu  
275 280 285

Gln Pro Phe Leu Leu Val Thr Tyr Leu Trp Gln Tyr Met Pro Thr Trp  
290 295 300

10 Ala Trp Trp Ile Thr Asn Lys Met Gly Lys Lys Arg Ile Glu Asn Phe  
305 310 315 320

Lys Ser Gly Val Asp Ala Asp Ser Ser Tyr Phe Lys Ile Phe Lys Thr  
15 325 330 335

Lys His Asp

20

(2) INFORMATION FOR SEQ ID NO: 249:

(i) SEQUENCE CHARACTERISTICS:  
25 (A) LENGTH: 96 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

30 Met Gly Ala Arg Pro Gly Gly His Pro Gln Lys Trp Ser Phe Leu Trp  
1 5 10 15

Ser Leu Ala Leu Trp Leu Pro Leu Ala Leu Ser Val Ser Leu Phe Leu  
20 25 30

35 Gly Leu Ser Leu Ser Pro Pro Gln Pro Gly Leu Ser Leu Trp Cys Thr  
35 40 45

40 Leu Ser Tyr Cys Cys Glu Gln Trp Lys Phe Lys Gly Thr Pro Ser Pro  
50 55 60

Ala Leu Leu Asn Leu Gly Thr Gln Pro Lys Lys Asp Lys Lys Leu Glu  
65 70 75 80

45 Asp Ser Ile Ala Thr Gln Leu Arg Xaa Leu Pro Glu Lys Asn Ser Asn  
85 90 95

50

(2) INFORMATION FOR SEQ ID NO: 250:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 79 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

60

Met Ala Leu Thr Phe Leu Leu Val Leu Leu Thr Leu Ala Thr Leu Cys  
 1 5 10 15

Thr Arg Leu His Arg Asn Phe Arg Arg Gly Glu Ser Ile Tyr Trp Gly  
 5 20 25 30

Pro Thr Ala Asp Ser Gln Asp Thr Val Ala Ala Val Leu Lys Arg Arg  
 35 40 45

10 Leu Leu Gln Pro Ser Arg Arg Val Lys Arg Ser Arg Arg Arg Pro Xaa  
 50 55 60

Xaa Pro Pro Thr Pro Asp Ser Gly Pro Glu Gly Glu Ser Ser Glu  
 65 70 75

15

## (2) INFORMATION FOR SEQ ID NO: 251:

20 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 354 amino acids  
     (B) TYPE: amino acid  
     (D) TOPOLOGY: linear  
     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

25 Met Gly Pro Ser Thr Pro Leu Leu Ile Leu Phe Leu Leu Ser Trp Ser  
 1 5 10 15

30 Gly Pro Leu Gln Gly Gln Gln His His Leu Val Glu Tyr Met Glu Arg  
 20 25 30

Arg Leu Ala Ala Leu Glu Glu Arg Leu Ala Gln Cys Gln Asp Gln Ser  
 35 40 45

35 Ser Arg His Ala Ala Glu Leu Arg Asp Phe Lys Asn Lys Met Leu Pro  
 50 55 60

Leu Leu Glu Val Ala Glu Lys Glu Arg Glu Ala Leu Arg Thr Glu Ala  
 65 70 75 80

40 Asp Thr Ile Ser Gly Arg Val Asp Arg Leu Glu Arg Glu Val Asp Tyr  
 85 90 95

45 Leu Glu Thr Gln Asn Pro Ala Leu Pro Cys Val Glu Phe Asp Glu Lys  
 100 105 110

Val Thr Gly Gly Pro Gly Thr Lys Gly Lys Gly Arg Arg Asn Glu Lys  
 115 120 125

50 Tyr Asp Met Val Thr Asp Cys Gly Tyr Thr Ile Ser Gln Val Arg Ser  
 130 135 140

Met Lys Ile Leu Lys Arg Phe Gly Gly Pro Ala Gly Leu Trp Thr Lys  
 145 150 155 160

55 Asp Pro Leu Gly Gln Thr Glu Lys Ile Tyr Val Leu Asp Gly Thr Gln  
 165 170 175

60 Asn Asp Thr Ala Phe Val Phe Pro Arg Leu Arg Asp Phe Thr Leu Ala  
 180 185 190

350

Met Ala Ala Arg Lys Ala Ser Arg Val Arg Val Pro Phe Pro Trp Val  
 195 200 205  
 5 Gly Thr Gly Gln Leu Val Tyr Gly Gly Phe Leu Tyr Phe Ala Arg Arg  
 210 215 220  
 Pro Pro Gly Arg Pro Gly Gly Glu Met Glu Asn Thr Leu Gln  
 225 230 235 240  
 10 Leu Ile Lys Phe His Leu Ala Asn Arg Thr Val Val Asp Ser Ser Val  
 245 250 255  
 Phe Pro Ala Glu Gly Leu Ile Pro Pro Tyr Gly Leu Thr Ala Asp Thr  
 15 260 265 270  
 Tyr Ile Asp Leu Ala Ala Asp Glu Glu Gly Leu Trp Ala Val Tyr Ala  
 275 280 285  
 20 Thr Arg Glu Asp Asp Arg His Leu Cys Leu Ala Lys Leu Asp Pro Gln  
 290 295 300  
 Thr Leu Asp Thr Glu Gln Gln Trp Asp Thr Pro Cys Pro Arg Glu Asn  
 305 310 315 320  
 25 Ala Glu Ala Ala Phe Xaa Ile Cys Gly Thr Leu Tyr Val Val Tyr Asn  
 325 330 335  
 30 Thr Arg Pro Ala Ser Arg Ala Arg Ile Gln Cys Ser Phe Asp Ala Ser  
 340 345 350  
 Gly Pro

35

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:  
 40 (A) LENGTH: 109 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:  
 45 Met Leu Cys Ile Asn Gly Thr Thr Pro Arg Pro Leu Pro Val Pro Ser  
 1 5 10 15  
 Pro Phe Gly Cys Met Ile Phe Phe Phe Phe Lys Asn Pro Trp Lys Gln  
 20 25 30  
 50 Arg Leu Leu Gln Gly Trp Leu Gly Ala Arg Pro Ile His Leu Leu Gly  
 35 40 45  
 Tyr Leu Pro Leu Ser Leu Leu Trp Cys Pro Phe Pro Leu Pro Cys Ala  
 55 50 55 60  
 Arg Cys Ser Val Val Tyr Ile Ser Ser Pro Arg His Gly Ala His Ala  
 65 70 75 80  
 60 Pro Arg Asp Met Ile Leu Ser Leu Val Leu Ala His Gly Ala Leu Tyr

85                    90                    95

Lys Glu Leu Gly Gly Arg Gly Arg Lys Trp Glu Pro Ser  
                       100                    105

5

## (2) INFORMATION FOR SEQ ID NO: 253:

## 10                    (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

15                    Met Phe Tyr Phe Leu Pro Leu Ile Phe Pro Ala Phe Pro Pro Trp Ala  
                       1                    5                    10                    15

20                    Phe Arg Leu Ser Thr Leu Phe Thr Ile Ile Ser Trp Ser Glu Asp Ser  
                       20                    25                    30

Asn Asn Ser Gln Val Tyr Met Asn Cys Val Cys Ser Phe  
                       35                    40                    45

25

## (2) INFORMATION FOR SEQ ID NO: 254:

## 30                    (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

35                    Met Ala Gly Gly Arg Cys Gly Pro Xaa Leu Thr Ala Leu Leu Ala Ala  
                       1                    5                    10                    15

Trp Ile Ala Ala Val Ala Ala Thr Ala Gly Pro Glu Glu Ala Ala Leu  
                       20                    25                    30

40                    Pro Pro Glu Gln Ser Arg Val Gln Pro Met Thr Ala Ser Asn Trp Thr  
                       35                    40                    45

45                    Leu Val Met Glu Gly Glu Trp Met Leu Lys Phe Tyr Ala Pro Trp Cys  
                       50                    55                    60

Pro Ser Cys Gln Gln Thr Asp Ser Glu Trp Glu Ala Phe Ala Lys Asn  
                       65                    70                    75                    80

50                    Gly Glu Ile Leu Gln Ile Ser Val Gly Lys Val Asp Val Ile Gln Glu  
                       85                    90                    95

Pro Gly Leu Ser Gly Arg Phe Phe Val Thr Thr Leu Pro Ala Phe Phe  
                       100                    105                    110

55                    His Ala Lys Asp Gly Ile Phe Arg Arg Tyr Arg Gly Pro Gly Ile Phe  
                       115                    120                    125

60                    Glu Asp Leu Gln Asn Tyr Ile Leu Glu Lys Lys Trp Gln Ser Val Glu  
                       130                    135                    140

Pro Leu Thr Gly Trp Lys Ser Pro Ala Ser Leu Thr Met Ser Gly Met  
 145 150 155 160

5 Ala Gly Leu Phe Ser Ile Ser Gly Lys Ile Trp His Leu His Asn Tyr  
 165 170 175

Phe Thr Val Thr Leu Gly Ile Pro Ala Trp Cys Ser Tyr Val Phe Phe  
 180 185 190

10 Val Ile Ala Thr Leu Val Phe Gly Leu Phe Met Gly Leu Val Leu Val  
 195 200 205

15 Val Ile Ser Glu Cys Phe Tyr Val Pro Leu Pro Arg His Leu Ser Glu  
 210 215 220

Arg Ser Glu Gln Asn Arg Arg Ser Glu Glu Ala His Arg Ala Glu Gln  
 225 230 235 240

20 Leu Gln Asp Ala Glu Glu Lys Asp Asp Ser Asn Glu Glu Asn  
 245 250 255

Lys Asp Ser Leu Val Asp Asp Glu Glu Lys Glu Asp Leu Gly Asp  
 260 265 270

25 Glu Asp Glu Ala Glu Glu Glu Glu Asp Asn Leu Ala Ala Gly  
 275 280 285

30 Val Asp Glu Glu Arg Ser Glu Ala Asn Asp Gln Gly Pro Pro Gly Glu  
 290 295 300

Asp Gly Val Thr Arg Glu Xaa Ser Arg Ala Xaa  
 305 310 315

35

(2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 53 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

45 Met Leu Lys Ala Leu Phe Arg Thr Leu Gln Ala Met Leu Leu Gly Val  
 1 5 10 15

Trp Ile Leu Leu Leu Ala Ser Leu Ala Pro Leu Trp Leu Tyr Cys  
 20 25 30

50 Trp Arg Met Phe Pro Thr Lys Gly Lys Arg Asp Gln Lys Glu Met Leu  
 35 40 45

Glu Val Ser Gly Ile  
 55 50

(2) INFORMATION FOR SEQ ID NO: 256:

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

Met	Ile	His	Leu	Gly	His	Ile	Leu	Phe	Leu	Leu	Leu	Leu	Pro	Val	Ala
1															
														15	

10	Ala	Ala	Gln	Thr	Thr	Pro	Gly	Glu	Arg	Ser	Ser	Leu	Pro	Ala	Phe	Tyr
														30		

15	Pro	Gly	Thr	Ser	Gly	Ser	Cys	Ser	Gly	Ser	Leu	Ser	Leu	Pro		
														45		

15	Leu	Leu	Ala	Gly	Leu	Val	Ala	Ala	Asp	Ala	Val	Ala	Ser	Leu	Leu	Ile
														50		

20	Val	Gly	Ala	Val	Phe	Leu	Cys	Ala	Arg	Pro	Arg	Arg	Ser	Pro	Ala	Gln
														65		

														85		

25

(2) INFORMATION FOR SEQ ID NO: 257:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

35	Pro	Gly	His	Leu	Leu	Pro	His	Lys	Trp	Glu	Asn	Cys				
														1	5	10

40 (2) INFORMATION FOR SEQ ID NO: 258:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1852 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

50	TGGCATCTGT	GAGCAGCTGC	CAGGCTCCGG	CCAGGATCCC	TTCCCTCTCC	TCATTGGCTG		60
	ATGGATCCCA	AGGGGCTCCT	CTCCTTGACC	TTCGTGCTGT	TTCTCTCCCT	GGCTTTGGG		120
55	GCAAGCTACG	GAACAGGTGG	GCGCATGATG	AACTGCCCAA	AGATTCTCCG	GCAGTTGGGA		180
	ACCAAAGTGC	TGCTGCCCT	GACATATGAA	AGGATAAATA	AGAGCATGAA	CAAAGCATC		240
	CACATTGTCG	TCACAATGGC	AAAATCACTG	GAGAACAGTG	TCGAGAACAA	AATAGTGTCT		300
60	CTTGATCCAT	CCGAAGCAGG	CCCTCACGT	TATCTAGGAG	ATCGCTACAA	GTTTATCTG		360

	GAGAATCTCA CCCTGGGGAT ACGGGAAAGC AGGAAGGAGG ATGAGGGATG GTACCTTATG	420
5	ACCTGGAGA AAAATGTTTC AGTTCAAGCGC TTTTGCCCTGC AGTTGAGGCT TTATGAGCAG	480
	GTCTCCACTC CAGAAATTAA AGTTTTAAC AAGACCCAGG AGAACGGGAC CTGCACCTTG	540
	ATACTGGCT GCACAGTGGA GAAGGGGGAC CATGTGGCTT ACAGCTGGAG TGAAAAGGCG	600
10	GGCACCCACC CACTGAACCC AGCCAACAGC TCCCACCTCC TGTCCCTCAC CCTCGGCCCC	660
	CAGCATGCTG ACAATATCTA CATCTGCACC GTGAGCAACC CTATCAGCAA CAATTCCCAG	720
15	ACCTTCAGCC CGTGGCCCGG ATGCAGGACA GACCCCTCAG AAACAAAACC ATGGGCAGTG	780
	TATGCTGGC TGTTAGGGGG TGTCACTCATG ATTCTCATCA TGTTGGTAAT ACTACAGTIG	840
	AGAAGAAGAG GTAAAACGAA CCATTACCAAG ACAACACTGG AAAAAAAAAG CCTTACGATC	900
20	TATGCCAAG TCCAGAAACC AGGTGACACT CATCATCAGA CTTGGACTT ATTCTAATCC	960
	AGGATGACCT TATTTGAAA TCCTTATCTT GACATCTGTG AAGACCTTTA TTCAAATAAA	1020
25	GTCACATTTT GACATTCTGC GAGGGCTGG AGCCGGGCCG GGGCGATGTG GAGCGCGGCC	1080
	CGCGCGGGGG CTGCCTGGCC GGTGCTGTIG GGGCTGCTGC TGGCGCTGTT AGTGCCTGGC	1140
	GGTGGTGCCG CCAAGACCGG TGCGGAGCTC GTGACTGCGG GTCGGTGCTG AAGCTGCTCA	1200
30	ATACGCACCA CGGGTGGCGC TGCACTCGCA CGACATCAA TACGGATCCG GCAGCGGCCA	1260
	GCAATCGGTG ACCGGCGTAG AGGTGGGAGC GACGAATAGC TACTGGCGGA TCCGGCGCGG	1320
	CTCGGAGGGG GGTGCCCGCG CGGGTCCCCG GTGCGCTGCG GGCAGGGCGT GAGGTCACAC	1380
35	ATGTGCTTAC GGGCAAGAAC CTGCACACGC ACCACTTCCC GTCGCCGCTG TCCAACAACC	1440
	AGGAAGTGAG TGCCAAAGGG GAAGACGGCG AGGGCGACGA CCTGGACCTA TGGACAGTGC	1500
40	GCTGCTCTGC TCTGGACAGC ACTGGGAGCG TGAGGCTGCT GTGGCGCTT CCAGCATGTG	1560
	GCACCTCTGT GGTCCTGTG AGTCACGGTA GCAGTATGGA AGCCCCATCC GTGGGCAGCA	1620
	TGAGGTCCAC GCATGCCAG TGCCAAACACG CACAATACGT GGAAGGCCAT GGAAGGCATC	1680
45	TTCATCAAGC CTAGTGTGGA GCCCTCTGCA GGTCACGATG AACTCTGAGT GTGTGGATGG	1740
	ATGGGTGGAT GGAGGGTGGC AGGTGGGCG TCTGCAGGGC CACTCTTGGC AGAGACTTTG	1800
50	GGTTTGTAGG GGTCCTCAAG TGCCTTGTG ATTAAAGAAT GTTGGTCTAT GA	1852

## 55 (2) INFORMATION FOR SEQ ID NO: 259:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 371 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu  
1 5 10 15

5 Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg  
20 25 30

10 Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp  
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu  
50 55 60

15 Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr  
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser  
85 90 95

20 Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu  
100 105 110

25 Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser  
115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr  
130 135 140

30 Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Lys Ile Arg Asn  
145 150 155 160

Lys Arg Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Val Tyr Val Gly  
165 170 175

35 Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu  
180 185 190

40 Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp  
195 200 205

Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr  
210 215 220

45 Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu  
225 230 235 240

Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro  
245 250 255

50 Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu  
260 265 270

55 Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys  
275 280 285

Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala  
290 295 300

60 Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser

356

305 310 315 320

Ala Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Ser Ser Glu Pro  
325 330 335

5

Leu Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys  
340 345 35010 Gly Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg  
355 360 365Tyr Ser Gly  
370

15

(2) INFORMATION FOR SEQ ID NO: 260:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

25

Cys Arg Cys Ala Ser Gly Phe Thr Gly Glu Asp Cys  
1 5 10

30

(2) INFORMATION FOR SEQ ID NO: 261:

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

40

Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys  
1 5 10

45

(2) INFORMATION FOR SEQ ID NO: 262:

50

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 12 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

55

Cys Leu Asn Leu Pro Gly Ser Tyr Gln Cys Gln Cys  
1 5 10

60

(2) INFORMATION FOR SEQ ID NO: 263:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid

(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

Cys Lys Cys Leu Thr Gly Phe Thr Gly Gln Lys Cys  
5 1 5 10

(2) INFORMATION FOR SEQ ID NO: 264:

10

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 12 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

Cys Gln Cys Leu Gln Gly Phe Thr Gly Gln Tyr Cys  
1 5 10

20

(2) INFORMATION FOR SEQ ID NO: 265:

25

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 127 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

30

Gly Leu Ala Cys Trp Leu Ala Gly Val Ile Phe Ile Asp Arg Lys Arg  
1 5 10 15

Thr Gly Asp Ala Ile Ser Val Met Ser Glu Val Ala Gln Thr Leu Leu  
20 25 30

35

Thr Gln Asp Val Xaa Val Trp Val Phe Pro Glu Gly Thr Arg Asn His  
35 40 45

40

Asn Gly Ser Met Leu Pro Phe Lys Arg Gly Ala Phe His Leu Ala Val  
50 55 60

Gln Ala Gln Val Pro Ile Val Pro Ile Val Met Ser Ser Tyr Gln Asp  
65 70 75 80

45

Phe Tyr Cys Lys Lys Glu Arg Arg Phe Thr Ser Gly Gln Cys Gln Val  
85 90 95

Arg Val Leu Pro Pro Val Pro Thr Glu Gly Leu Thr Pro Asp Asp Val  
100 105 110

50

Pro Ala Leu Ala Asp Arg Val Arg His Ser Met Leu His Cys Phe  
115 120 125

55

(2) INFORMATION FOR SEQ ID NO: 266:

60

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 98 amino acids  
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

5 Pro Ser Ala Lys Tyr Phe Phe Lys Met Ala Phe Tyr Asn Gly Trp Ile  
1 5 10 15

Leu Phe Leu Ala Val Leu Ala Ile Pro Val Cys Ala Val Arg Gly Arg  
20 25 30

10 Asn Val Glu Asn Met Lys Ile Leu Arg Leu Met Leu Leu His Ile Lys  
35 40 45

Tyr Leu Tyr Gly Ile Arg Val Glu Val Arg Gly Ala His His Phe Pro  
50 55 60

15 Pro Ser Gln Pro Tyr Val Val Val Ser Asn His Gln Ser Ser Leu Asp  
65 70 75 80

Leu Leu Gly Met Met Glu Val Leu Pro Gly Arg Cys Val Pro Ile Ala  
20 85 90 95

Lys Arg

25

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 9 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

35 Thr Val Phe Arg Glu Ile Ser Thr Asp  
1 5

40 (2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 11 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

Leu Trp Ala Gly Ser Ala Gly Trp Pro Ala Gly  
1 5 10

50

(2) INFORMATION FOR SEQ ID NO: 269:

55 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

60

Ser Ile Leu Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala  
1 5 10 15

5 Ser Lys His Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg  
20 25

10 (2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

Met Ala Tyr His Gly Leu Thr Val  
1 5

20

(2) INFORMATION FOR SEQ ID NO: 271:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

30 Ile Ser Ala Ala Arg Val  
1 5

35 (2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

Pro Asp Val Ser Glu Phe Met Thr Arg Leu Phe  
1 5 10

45

(2) INFORMATION FOR SEQ ID NO: 273:

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

55

Phe Asp Pro Val Arg Val Asp Ile Thr Ser Lys Gly Lys Met Arg Ala  
1 5 10 15

Arg

360

Applicant's or agent's file reference number	PS001PCT	International application	Unassigned
--	----------	---------------------------	------------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

**A.** The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution  
American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97901

**C. ADDITIONAL INDICATIONS** *(leave blank if not applicable)* This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** *(if the indications are not for all designated States)*

**E. SEPARATE FURNISHING OF INDICATIONS** *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

361

Applicant's or agent's file reference number	PS001PCT	International applicatio.	Unassigned
--	----------	---------------------------	------------

## **INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> . line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit      February 26, 1997	Accession Number      97898
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )	
The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

**For receiving Office use only**

This sheet was received with the international application

**Authorized officer**

**For International Bureau use only**

This sheet was received by the International Bureau on:

**Authorized officer**

362

Applicant's or agent's file reference number	1S001PCT	International application	Unassigned
--	----------	---------------------------	------------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit    May 15, 1997	Accession Number    209044
C. ADDITIONAL INDICATIONS ( <i>leave blank if not applicable</i> )    This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE ( <i>if the indications are not for all designated States</i> )	
E. SEPARATE FURNISHING OF INDICATIONS ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

363

Applicant's or agent's file reference number	7S001PCT	International application	Unassigned	144562
--	----------	---------------------------	------------	--------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 64, line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution American Type Culture Collection

Address of depositary institution (including postal code and country)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97899

**C. ADDITIONAL INDICATIONS** (leave blank if not applicable) This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (if the indications are not for all designated States)

**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only



This sheet was received with the international application

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer



Authorized officer

364

Applicant's or agent's file reference number	PS001PCT	International application	Unassigned	774482
--	----------	---------------------------	------------	--------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>65</u> line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution <u>American Type Culture Collection</u>  Address of depositary institution ( <i>including postal code and country</i> )  <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209045</u>
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

365		
Applicant's or agent's file reference number	PS001PCT	International application
		Unassigned
		17-25

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

**A.** The indications made below relate to the microorganism referred to in the description on page 64, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution

American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97900

**C. ADDITIONAL INDICATIONS** *(Leave blank if not applicable)* This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** *(If the indications are not for all designated States)*

**E. SEPARATE FURNISHING OF INDICATIONS** *(Leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

366

Applicant's or agent's file reference number	'S001PCT	International application	Unassigned	111111111111111111111111
--	----------	---------------------------	------------	--------------------------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description on page 64, line N/A</b>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209046
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only		For International Bureau use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:		
Authorized officer 	Authorized officer		

367

Applicant's or agent's file reference number	'S001PCT	International application	Unassigned	1 2 3 4 5 6
--	----------	---------------------------	------------	-------------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65 . line N/A	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit April 28, 1997	Accession Number 209010
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	
For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer 	
For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on _____	
Authorized officer	

368

Applicant's or agent's file reference number	I5001PCT	International application	Unassigned	SEARCHED
--	----------	---------------------------	------------	----------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 65 . line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution      American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 29, 1997      Accession Number 209085

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)      This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

369

Applicant's or agent's file reference number	'S001PCT	International application	Unassigned	SEARCHED
--	----------	---------------------------	------------	----------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65 line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution      American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997      Accession Number 97897

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)      This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

370

Applicant's or agent's file reference number	PS001PCT	International application	<input type="checkbox"/> Unassigned	SEARCHED
--	----------	---------------------------	-------------------------------------	----------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 65, line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution  
American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209043

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

371

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	162 222
--	---------	---------------------------	------------	---------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>73</u> . line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit      September 4, 1997	Accession Number      209236
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	
For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer 	
For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

372		
Applicant's or agent's file reference number	S001PCT	International application
		Unassigned <span style="float: right;">四千六百</span>

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description on page <u>73</u>, line <u>N/A</u></b>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit    May 29, 1997	Accession Number    209084
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only

<input checked="" type="checkbox"/> This sheet was received with the international application
Authorized officer 

For International Bureau use only

<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer

373

Applicant's or agent's file reference number	S001PCT	International application	Unassigned
--	---------	---------------------------	------------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 76 . line N/A	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209048
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only

This sheet was received with the international application

Authorized officer 

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

374

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	24452
--	---------	---------------------------	------------	-------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

**A.** The indications made below relate to the microorganism referred to in the description  
on page 76 line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution

American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97902

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*)

This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

375

Applicant's or agent's file reference number	S001PCT	International applicant	Unassigned	SEARCHED
--	---------	-------------------------	------------	----------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description on page 77</b>		<b>. line N/A</b>
<b>B. IDENTIFICATION OF DEPOSIT</b>		
Name of depositary institution      American Type Culture Collection		
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America		
Date of deposit	February 26, 1997	Accession Number      97903
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input type="checkbox"/>		
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )		
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )		
For receiving Office use only		For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application		<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer		Authorized officer

376

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	Office
--	---------	---------------------------	------------	--------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 77 line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution  
American Type Culture Collection

Address of depositary institution (including postal code and country)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997 Accession Number 209049

C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)****E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)**

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

377

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	Office
--	---------	---------------------------	------------	--------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 80, line N/A

**B. IDENTIFICATION OF DEPOSIT**Further deposits are identified on an additional sheet 

Name of depositary institution American Type Culture Collection

Address of depositary institution (including postal code and country)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997 Accession Number 97904

C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

378

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	UNPCT
--	---------	---------------------------	------------	-------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 80 line N/A	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209050
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	
For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application  Authorized officer	
For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:  Authorized officer	

379

Applicant's or agent's file reference number	S001PCT	International application	Unassigned	144-22
--	---------	---------------------------	------------	--------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A. The indications made below relate to the microorganism referred to in the description on page 82.</b> line N/A	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> ) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit April 4, 1997	Accession Number 97976
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> ) This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> ) The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

380

Applicant's or agent's file reference number	3001PCT	International application	Unassigned	SEARCHED
--	---------	---------------------------	------------	----------

**INDICATIONS RELATING TO A DEPOSITED MICROORGANISM**

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit    May 15, 1997	Accession Number    209047
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )    This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

**What Is Claimed Is:**

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
  - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
  - (f) a polynucleotide which is a variant of SEQ ID NO:X;
  - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
  - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
  - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

10

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

15

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

20

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

30

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in

35 ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

- (g) a variant of SEQ ID NO:Y;
- (h) an allelic variant of SEQ ID NO:Y; or
- (i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the

5 full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of  
claim 11.

10 14. A recombinant host cell that expresses the isolated polypeptide of claim  
11.

15. A method of making an isolated polypeptide comprising:

15 (a) culturing the recombinant host cell of claim 14 under conditions such that  
said polypeptide is expressed; and  
(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

20 17. A method for preventing, treating, or ameliorating a medical condition,  
comprising administering to a mammalian subject a therapeutically effective amount of  
the polypeptide of claim 11 or the polynucleotide of claim 1.

25 18. A method of diagnosing a pathological condition or a susceptibility to a  
pathological condition in a subject comprising:  
(a) determining the presence or absence of a mutation in the polynucleotide of  
claim 1; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological  
30 condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a  
pathological condition in a subject comprising:

35 (a) determining the presence or amount of expression of the polypeptide of  
claim 11 in a biological sample; and  
(b) diagnosing a pathological condition or a susceptibility to a pathological  
condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

5 (b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

10 22. A method of identifying an activity in a biological assay, wherein the method comprises:

(a) expressing SEQ ID NO:X in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

15 (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.

Applicant's or agent's file  
reference number

PS001PCT

International applicatio

Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97900

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application? <input checked="" type="checkbox"/> Unassigned
--	----------	---

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>65</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit    May 15, 1997	Accession Number    209043
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )    This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209044

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

### **UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### **DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### **SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### **NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

- A. The indications made below relate to the microorganism referred to in the description  
on page 65, line N/A

## B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet 

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209045

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet 

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209046 ..

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

### UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International applicatio

to. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209047

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application to Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 76, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209048 ..

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application	J. Unassigned
--	----------	---------------------------	---------------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>77</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit    May 15, 1997	Accession Number    209049
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input checked="" type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )  -----	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

### UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

P5001PCT

International application

Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 80, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 15, 1997

Accession Number 209050

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

### UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No. unassigned
--	----------	--

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 73, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution      American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit    September 4, 1997

Accession Number    209236 ..

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)    This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer:

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

### UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International applicat

No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 65, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit April 28, 1997

Accession Number 209010

### C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)

This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application number	10. Unassigned
--	----------	----------------------------------	----------------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 29, 1997

Accession Number 209085

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g. "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application	Io. Unassigned
--	----------	---------------------------	----------------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b>	
Name of depositary institution      American Type Culture Collection	
Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Address of depositary institution ( <i>including postal code and country</i> )  12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit      February 26, 1997	Accession Number      97901 ..
<b>C. ADDITIONAL INDICATIONS</b> ( <i>leave blank if not applicable</i> )      This information is continued on an additional sheet <input type="checkbox"/>	
in respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> ( <i>if the indications are not for all designated States</i> )  -----	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> ( <i>leave blank if not applicable</i> )  The indications listed below will be submitted to the International Bureau later ( <i>specify the general nature of the indications, e.g., "Accession Number of Deposit"</i> )	

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application to Unassigned
--	----------	---

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

**A.** The indications made below relate to the microorganism referred to in the description on page 77, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997 Accession Number 97903

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

## **CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## **NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## **AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## **FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International applicatio

No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97898

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

## UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

## DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

## SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

## NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 80, line N/A

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97904

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer



Authorized officer

## **CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## **NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## **AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## **FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application to. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 73, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit May 29, 1997

Accession Number 209084 ..

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*If the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

## UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

## DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

## SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

## NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application to. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 64, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97899

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

### UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

### DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

### SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

### NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

PS001PCT

International application No. Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 65, line N/A

### B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97897

### C. ADDITIONAL INDICATIONS (*leave blank if not applicable*)

This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

### D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

### E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer



For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

## UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

## DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

## SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

## NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application	Unassigned
--	----------	---------------------------	------------

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

**A.** The indications made below relate to the microorganism referred to in the description  
on page 82, line N/A

**B. IDENTIFICATION OF DEPOSIT**

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit April 4, 1997

Accession Number 97976 "

**C. ADDITIONAL INDICATIONS** (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

**D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE** (*if the indications are not for all designated States*)

**E. SEPARATE FURNISHING OF INDICATIONS** (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**Page 2**

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

## CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

## NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

## AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

## FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

**NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file  
reference number

P5001PCT

International application

Unassigned

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 76, line N/A

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet

Name of depositary institution American Type Culture Collection

Address of depositary institution (*including postal code and country*)

12301 Parklawn Drive  
Rockville, Maryland 20852  
United States of America

Date of deposit February 26, 1997

Accession Number 97902

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only



This sheet was received with the international application

Authorized officer

For International Bureau use only



This sheet was received by the International Bureau on:

Authorized officer



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A3	(11) International Publication Number: WO 98/39446 (43) International Publication Date: 11 September 1998 (11.09.98)		
C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17		MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).		
(21) International Application Number:	PCT/US98/04482			
(22) International Filing Date:	6 March 1998 (06.03.98)			
(30) Priority Data:				
60/040,162	7 March 1997 (07.03.97)	US		
60/040,333	7 March 1997 (07.03.97)	US		
60/038,621	7 March 1997 (07.03.97)	US		
60/040,161	7 March 1997 (07.03.97)	US		
60/040,626	7 March 1997 (07.03.97)	US		
60/040,334	7 March 1997 (07.03.97)	US		
60/040,336	7 March 1997 (07.03.97)	US		
60/040,163	7 March 1997 (07.03.97)	US		
60/043,580	11 April 1997 (11.04.97)	US		
60/043,568	11 April 1997 (11.04.97)	US		
<i>(Continued on the following page)</i>				
(71) Applicant (for all designated States except US):	HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).			
(72) Inventors; and				
(75) Inventors/Applicants (for US only):	RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BEDNARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Damestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda,			
(81) Designated States:	AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW. ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).			
Published				
With international search report.				
With an indication in relation to a deposited microorganism furnished under Rule 13 <sup>bis</sup> separately from the description.				
Date of receipt by the International Bureau:				
06 April 1998 (06.04.98)				
(88) Date of publication of the international search report:	23 December 1998 (23.12.98)			

(54) Title: 70 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

60/043,314	11 April 1997 (11.04.97)	US	60/047,598	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US
60/043,569	11 April 1997 (11.04.97)	US	60/047,613	23 May 1997 (23.05.97)	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,311	11 April 1997 (11.04.97)	US	60/047,582	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,671	11 April 1997 (11.04.97)	US	60/047,596	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,674	11 April 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,669	11 April 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,312	11 April 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05.97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,313	11 April 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05.97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,672	11 April 1997 (11.04.97)	US	60/047,599	23 May 1997 (23.05.97)	US	60/056,636	22 August 1997 (22.08.97)	US
60/043,315	11 April 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05.97)	US	60/056,874	22 August 1997 (22.08.97)	US
60/043,578	11 April 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05.97)	US	60/056,910	22 August 1997 (22.08.97)	US
60/043,576	11 April 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08.97)	US
60/043,670	11 April 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05.97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,594	23 May 1997 (23.05.97)	US	60/056,845	22 August 1997 (22.08.97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,589	23 May 1997 (23.05.97)	US	60/056,892	22 August 1997 (22.08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05.97)	US	60/056,632	22 August 1997 (22.08.97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,876	22 August 1997 (22.08.97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,881	22 August 1997 (22.08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,909	22 August 1997 (22.08.97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,875	22 August 1997 (22.08.97)	US
60/047,503	23 May 1997 (23.05.97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,862	22 August 1997 (22.08.97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22.08.97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22.08.97)	US
60/047,584	23 May 1997 (23.05.97)	US	60/056,630	22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
60/047,500	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
60/047,587	23 May 1997 (23.05.97)	US	60/056,662	22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047,492	23 May 1997 (23.05.97)	US	60/056,872	22 August 1997 (22.08.97)	US			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 98/04482

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6	C12N15/12	C12N5/10	C12N1/21	C07K14/47	C07K16/18
	C12Q1/68	G01N33/50	G01N33/53	G01N33/68	A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	L. HILLIER ET AL.: "The WashU-Merck EST Project 1997" EMBL SEQUENCE DATABASE, 6 March 1997, HEIDELBERG, FRG, XP002068123 zr78g10.r1 Soares NhHMPu S1 Homo sapiens cDNA clone 669570 5' similar to SW:FUCO_RAT P17164 Alpha-L-fucosidase precursor; Accession. Accession no. AA234924; --- -/-	1-3, 7-10,21

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

16 June 1998

Date of mailing of the international search report

16. 09. 1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patenttaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 98/04482

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 15 December 1996, HEIDELBERG, FRG, XP002068124 z140b11.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 504381 5' similar to TR:G182779 Lysosomal Enzyme Alpha-L-Fucosidase Accession no. AA151194 ---	1-3, 7-10,21
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 4 June 1996, HEIDELBERG, FRG, XP002068125 zc54a02.r1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 326090 5' similar to SW:FUCO_HUMAN P4066 tissue Alpha-L-Fucosidase precursor; Accession no. W52490 ---	1-3, 7-10,21
A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document ---	1-23
A	WO 97 04097 A (GENETICS INST) 6 February 1997 see the whole document ---	1-23
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 see the whole document ---	1-23
A	JACOBS K ET AL: "A novel method for isolating eukaryotic cDNA clones encoding secreted proteins." KEYSTONE SYMPOSIUM ON DENDRITIC CELLS: ANTIGEN PRESENTING CELLS OF T AND B LYMPHOCYTES, TAOS, NEW MEXICO, USA, MARCH 10-16, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (21A). 1995. 19. ISSN: 0733-1959, XP002027246 abstract no. C1-207 see abstract ---	1-23
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document ---	1-23
A	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document ---	1-23
		-/-

## INTERNATIO SEARCH REPORT

International Application No

PCT/US 98/04482

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	T. OCCHIODORO ET AL.: "Human alpha-L-Fucosidase: Complete coding sequence from cDNA clones" BIOCHEM. AND BIOPHYS. RES. COMMUNICATIONS, vol. 164, no. 1, 16 October 1989, ACADEMIC PRESS, NEW YORK, US, pages 439-445, XP002068126 cited in the application see the whole document -----	1-23

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 98/04482

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  
**Remark:** Although claim 17 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2.  Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see further information sheet

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1. Claims: (1-23) partially

-An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence consisting of SEQ ID no. 11; wherein said polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein encoding the sequence of SEQ ID no. 134 or the polypeptide encoded by the cDNA sequence included in ATCC Deposit no: HGCM20, which is hybridizable to SEQ ID no.11; a recombinant vector comprising said isolated nucleic acid molecule; a method of making a recombinant host cell comprising said isolated nucleic acid molecule; a recombinant host cell comprising said vector; an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence consisting of SEQ ID no. 134; an isolated antibody that binds specifically to said isolated polypeptide; a recombinant host cell that expresses said isolated polypeptide; a method of making said polypeptide; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said polypeptide; a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject using said polynucleotide and/or polypeptide sequences; a method for identifying a binding partner to said polypeptide; a gene corresponding to the cDNA sequences of SEQ ID no.11; a method for identifying an activity in a biological assay, by using the expression of SEQ ID no. 134;

Inventions 2 to 70. Claims: (1-23) partially

-Idem as subject 1 but limited to gene nos. 2 to 70 respectively cDNA clone sequences HLDBG33 to HMCAB89. (Invention 2 is limited to SEQ ID nos.12,81,135, and 204; Invention 3 is limited to SEQ ID nos.13 and 136; .....; Invention 70 is limited to SEQ ID nos.80 and 203;)

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/04482

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9707198 A	27-02-97	US 5707829 A AU 6712396 A AU 6768596 A EP 0839196 A EP 0851875 A WO 9704097 A	13-01-98 18-02-97 12-03-97 06-05-98 08-07-98 06-02-97
WO 9704097 A	06-02-97	US 5707829 A AU 6712396 A EP 0839196 A AU 6768596 A EP 0851875 A WO 9707198 A	13-01-98 18-02-97 06-05-98 12-03-97 08-07-98 27-02-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98
WO 9014432 A	29-11-90	US 5580753 A AT 147436 T AU 637620 B AU 5928990 A CA 2056997 A DE 69029657 D DK 473724 T EP 0473724 A ES 2099096 T JP 4506006 T US 5734037 A US 5414071 A	03-12-96 15-01-97 03-06-93 18-12-90 24-11-90 20-02-97 14-04-97 11-03-92 16-05-97 22-10-92 31-03-98 09-05-95
WO 9617925 A	13-06-96	AU 4639396 A CA 2206488 A FI 972390 A NO 972455 A	26-06-96 13-06-96 05-06-97 06-08-97



